The Dissertation Committee for Gurvani Bhupindersingh Certifies that this is the approved version of the following Dissertation:

MMTV Encodes a Protein Antagonist of Multiple Apobec Family Members

Committee:

Jaquelin P. Dudley, Supervisor

Haley Tucker

Christopher Sullivan

Karen Vasquez

Lauren Ehrlich

MMTV Encodes a Protein Antagonist of Multiple Apobec Family Members

by

Gurvani Bhupindersingh

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2018

Dedication

To my dearest parents for their unconditional love, support and belief in my dreams & to my husband for being my best friend.

Acknowledgements

My development as a scientist has been significantly influenced by my advisor Dr. Jaquelin Dudley. I would like to express my gratitude to her for providing me with the opportunity to be a part of her laboratory. Her enthusiasm and dedication to science has been very inspiring. I sincerely thank her for constant encouragement, guidance and support throughout my graduate studies.

I would like to thank Drs. Haley Tucker, Christopher Sullivan, Karen Vasquez and Lauren Ehrlich for graciously accepting to be a part of my dissertation committee and for providing constructive and valuable feedback on my research project.

I would like to express my deepest appreciation for all the past and current members of the Dudley Lab – Alex Aleman, Hyewon Byun, Eugenia Chen, Poulami Das, Yongqiang Gou, Mary Lozano, Frank Medina and Wendy Kaichun Xu. I would like to especially thank Mary Lozano for her work in mouse colony maintenance and breeding, which hugely contributed towards my experimental results. I would also like to acknowledge Sullivan and Ehrlich labs for sharing their reagents and protocols.

I am tremendously grateful to my parents for believing in my capabilities, providing me with the resources to fulfill my dreams and trusting my decisions.

Finally, I would like to thank my husband for his unwavering love, support and motivation, and my in-laws for their affection.

Abstract

MMTV Encodes a Protein Antagonist of Multiple Apobec Family Members

Gurvani Bhupindersingh, PhD.

The University of Texas at Austin, 2018

Supervisor: Jaquelin P. Dudley

Mouse Mammary Tumor Virus (MMTV) is a member of the genus *Betaretrovirus* in the family *Retroviridae*. MMTV, which primarily induces mammary carcinomas, has been extensively used as a model system to study human breast cancers. Due to its complex organization similar to human retroviruses, such as human immunodeficiency virus type 1(HIV-1) [12], studies of MMTV pathogenesis greatly aid in understanding viral interactions with the natural host immune system.

MMTV encodes a 33 kDa regulatory protein, Rem, which is synthesized from a doubly spliced form of viral genomic RNA [23, 48]. The precursor Rem is targeted to the endoplasmic reticulum (ER) for translation and is cleaved by signal peptidase into an N-terminal Signal Peptide (SP) and a C-terminal protein (Rem-CT) [50]. Previous studies have demonstrated that SP (also cleaved from envelope precursor protein) can manipulate the ER-associated degradation pathway (ERAD) and retrotranslocate to the cytoplasm

prior to nuclear entry to facilitate export of MMTV RNA [24]. SP thus has an analogous function to HIV-encoded Rev protein. The role of Rem-CT remains unknown.

Cytidine deaminases belonging to the apolipoprotein B messenger RNA editing catalytic polypeptide (APOBEC/Apobec) family are restriction factors for human and mouse retroviruses, respectively, as well as for retrotransposons and some DNA viruses [129, 132, 141-150, 162-166]. Murine Apobec3 (mA3)-deficient mice are more susceptible to MMTV infection [179], yet no known MMTV-encoded antagonist of mA3 has been identified. MMTV requires replication through B and T lymphocytes [33, 34, 37], both of which express Apobec family proteins, prior to mammary gland infection and tumorigenesis [93, 97].

Experiments were performed to address the role of Rem C-terminal sequences to antagonize Apobec enzymes and facilitate virus replication *in vivo.* Previous results have shown that mice infected with Rem-null MMTV developed mammary tumors with a lower incidence and increased latency compared to those infected with wild-type MMTV. Absence of Rem during *in vivo* infection resulted in lower proviral loads and increased mutations of the viral genome in mammary tumors. These differences were abolished in mice deficient for Activation-Induced Cytidine Deaminase (AID), a primordial member of the Apobec family that plays a role in shaping the adaptive immune response and development of human B-cell lymphomas. Tissue culture experiments revealed that Rem expression targets AID for proteasomal degradation. This study provides evidence that Rem is an HIV-Vif like protein and the first protein inhibitor of AID. In addition, these results suggest that Rem may antagonize other Apobec enzymes.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Mouse Mammary Tumor Virus (MMTV)

1.1.1 MMTV history and classification

MMTV is a complex murine retrovirus that was first reported in 1933 by Jackson Memorial Laboratory as an extrachromosomal factor associated with mammary tumors in inbred mice [1]. Later in 1936, Bittner demonstrated that this extra-chromosomal factor was transmitted through maternal milk, a virus then known as the "Bittner agent" [2]. Subsequently, MMTV isolated from milk of a high-mammary-tumor incidence mouse strain was shown to induce mammary tumors upon inoculation into a low-mammary-tumor incidence strain. This virus was also shown to have a reverse transcriptase (RT) [3, 4, 5]. Thus, MMTV was the first identified mammalian retrovirus that induced carcinomas. Since then MMTV has been used extensively as a model system to study human breast cancer [6].

MMTV is classified as a member of the *Retroviridae* family. Viral particles are enveloped, spherical and contain a dimer of single-stranded positive-sense RNA that exists as a helical ribonucleoprotein (RNP) encapsidated by the nucleocapsid protein (NC) along with reverse transcriptase (RT) and integrase (IN). The virus also has an icosahedral capsid composed of a single protein (CA), which surrounds RNP. The icosahedral capsid has an envelope consisting of a cellular phospholipid bilayer substituted with a viral envelope protein (Env) composed of a surface (SU) subunit and a transmembrane (TM) subunit. This

retrovirus exhibits a “B-type” morphology with prominent surface spikes, condensed core and has a characteristic particle size of 80-100 nm in diameter. Therefore, MMTV is the prototype for the genus *Betaretrovirus* based on its morphology [7, 8].

1.1.2 MMTV genome organization

MMTV genome is a single-stranded, positive-sense RNA of approximately 9 kb in length. Both ends of the genomic RNA have a 15 bp direct repeat region (R). Internal to the R region, unique sequence regions of 120 bp (U5) and 1,200 bp (U3) are found at the 5’ and 3’ ends, respectively. After infection, the MMTV RNA undergoes reverse transcription to yield a double-stranded viral DNA (provirus) that has long terminal repeats (LTRs) at both ends with the composition of U3-R-U5. The MMTV LTR contains hormone response elements (HREs), a mammary gland enhancer (MGE) and negative regulatory element (NRE), which determine hormone-response and tissue-specific transcription of the virus [9, 10, 11].

Unlike simple murine retroviruses such as murine leukemia virus (MLV), MMTV has a complex genome organization that is similar to human retroviruses like human immunodeficiency virus (HIV) and human T-cell leukemia virus (HTLV) [12]. The MMTV RNA encodes the following viral genes from 5’ to 3’: *gag* (group specific antigen), *dut* (dUTPase), *pro* (protease), *pol* (reverse transcriptase and integrase), *env* (envelope), *rem* (regulator of export/expression of MMTV mRNAs) and *sag* (superantigen). At least five transcripts are generated from integrated MMTV DNA (provirus) (Fig. 1.1).

Full-length unspliced MMTV RNA is transcribed from the U3 promoter in the 5' LTR. This RNA is translated into Gag, Pro, Pol proteins or is packaged into virions as the viral genome (Fig. 1.2). Using a double ribosomal frameshift during genomic RNA translation, three polyprotein precursors, Gag (Pr77), Gag-Pro (Pr110), and Gag-Pro-Pol (Pr160), are generated. Gag is processed by the viral protease PR (encoded by *pro*) into the capsid (CA), nucleocapsid (NC) and matrix (MA) proteins, as well as other proteins of unknown function [13, 14, 15]. The Gag-Pro (Pr110) precursor is cleaved to yield the viral protease (PR) and the dUTPase; the role of the accessory protein, dUTPase in MMTV pathogenesis in not known. However, in retroviruses, such as the equine infectious anemia virus (EIAV), the dUTPase facilitates virus replication in non-dividing cells by maintaining an adequate nucleotide pool and excluding dU from proviral DNA [16]. Thus, the MMTV dUTPase activity could potentially aid in viral replication in non-dividing cells. The Gag-Pro-Pol (Pr160) precursor is cleaved into RT (reverse transcriptase) and IN (integrase), which are needed to generate double-stranded proviral DNA and integration into the host chromosome, respectively [12]. A singly spliced mRNA is produced for expression of the envelope (*env*) gene, which encodes the Env precursor protein. Env is cleaved into surface (SU) and transmembrane (TM) domains by the host furin enzymes [12]. The extracellular SU domain has the receptor binding site (RBS) required to bind transferrin receptor-1 (TfR1) on the cells, whereas the TM domain mediates viral entry by allowing fusion with the cell membrane during receptor-mediated endocytosis [12, 17].

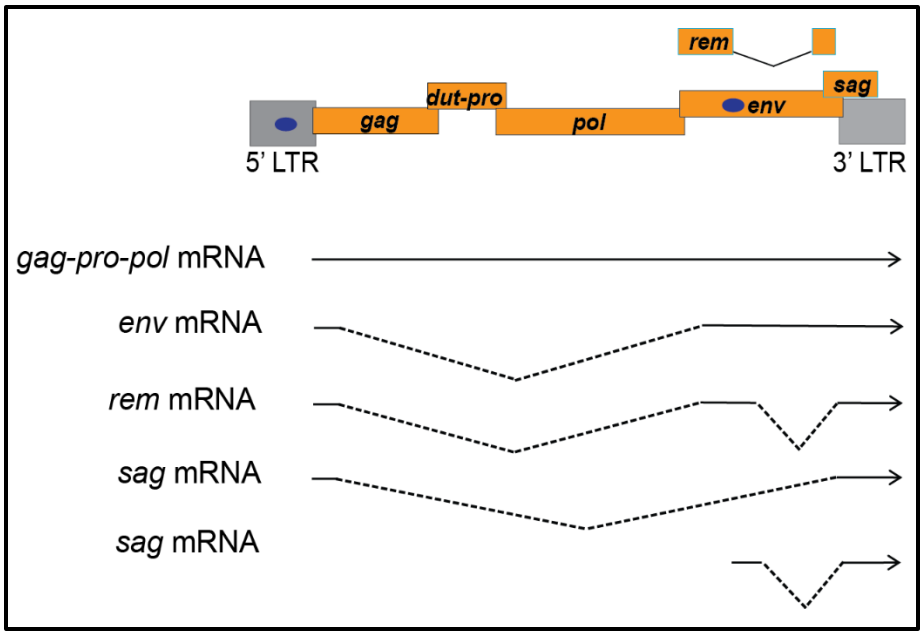

Figure 1.1: Diagram of the MMTV proviral genome and viral transcripts.

The boxed regions at the 5' and 3' ends represent the LTRs. Viral promoters are shown in blue circles. Four transcripts are generated from the LTR promoter: unspliced full-length genomic RNA, which may be exported to the cytosol for translation into Gag, Gag-Pro and Gag-Pro-Pol precursors. Alternatively, full-length RNA is processed into singly or doubly spliced mRNAs encoding Env, Sag or Rem. Sag may be translated from a singly spliced transcript generated from the 5' LTR or the internal envelope promoter.

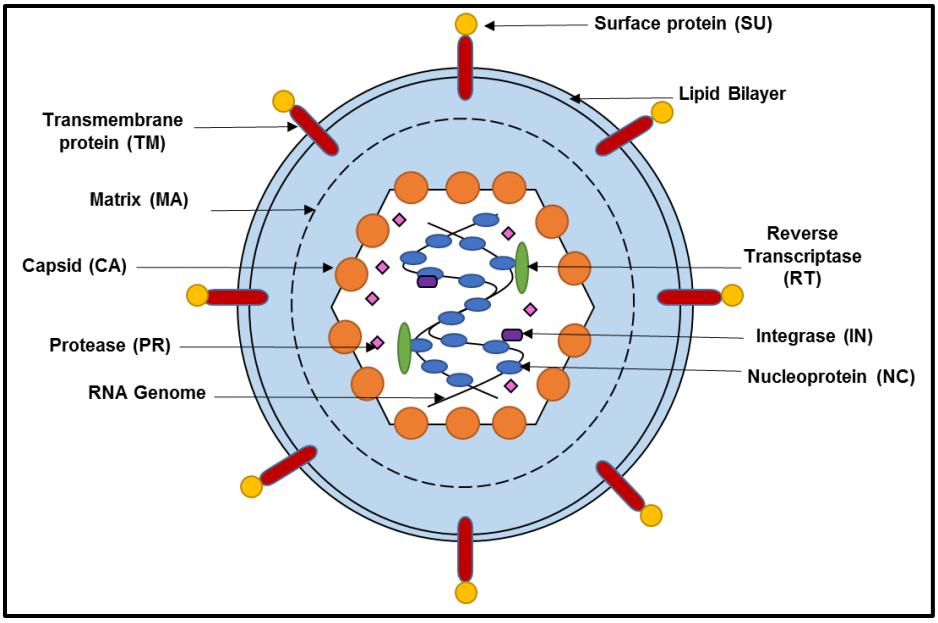

Figure 1.2: MMTV virion morphology.
A description of viral proteins and their functions is included in the text.

In addition, the MMTV genome encodes for another accessory protein Sag (superantigen) and a regulatory protein Rem (regulator of export/expression of MMTV mRNA). Singly spliced *sag* mRNA is produced from the upstream U3 promoter as well as the intragenic *env* promoter [18, 19]. Sag expression is critical for transmission of the virus from lymphoid cells to the mammary gland [12, 20, 21]. Rem is translated from a doubly spliced version of the *env* gene. Rem is a precursor protein that is cleaved into a functional signal peptide (SP) and a C-terminal product of unknown function. SP facilitates export of unspliced MMTV mRNA from the nucleus to the cytosol as well as some post-export activities and thus functions analogous to HIV-encoded Rev protein [22, 23, 24].

1.1.3 MMTV transmission and tissue tropism

Horizontal milk-borne transfer from infected mothers to pups (exogenous virus) is the most common route of MMTV infection. The virus can also be inherited vertically after exogenous MMTV infection of germline cells (endogenous proviruses or *Mtv*s) [25, 26]. At least 10 different exogenous and more than 30 endogenous *Mtv*s have been identified [27, 28]. Most of the commonly used inbred mouse strains have two to eight endogenous MMTV-related proviruses, such as *Mtv-6*, *8* and *9* in BALB/cJ mice and *Mtv-8*, *9* and *17* in C57BL/6 mice [29]. Most endogenous proviruses are defective in one or more genes and do not produce infectious virus. However, many *Mtv*s encode functional Sag, which induces intrathymic or peripheral deletion of Sag-reactive T cells and alters the mouse T-cell repertoire [30, 31]. Proteins encoded by endogenous *Mtv*s may serve as self-antigens to generate partial tolerance against MMTV. Strains that lack *Mtv*s often show resistance to MMTV transmission and tumorigenesis [32]. BALB/cJ congenic mouse strain lacking all endogenous *Mtv*s was resistant to MMTV infection compared to wild-type BALB/cJ mice, by oral as well as intraperitoneal modes of infection. Thus, the presence of endogenous *Mtv*s potentially aids infection by exogenous MMTVs. Alternatively, specific endogenous *Mtv*s may function to protect against exogenous MMTV infection. Transgenic mice expressing C3H MMTV ORF protein in addition to their endogenous *Mtv*s, were resistant to exogenous infection by the C3H MMTV strain [25].

MMTV acquired by newborn mice passes through the stomach and reaches the small intestine, where it first infects dendritic cells (DCs) and B cells within Peyer's patches [33, 34] (Fig. 1.3). Following infection, the viral genome undergoes reverse

transcription, integration and transcription that leads to expression of the Sag protein. Major histocompatibility complex (MHC) class II proteins present Sag on the surface of infected antigen-presenting cells (APCs). Sag-presenting cells interact with CD4+ T cells expressing specific T-cell receptor (TCR) Vβ chains [35, 36]. Sag presentation initiates a helper T-cell response by priming Sag-reactive T lymphocytes, resulting in their rapid proliferation and release of cytokines. Cytokine release leads to a pool of actively dividing lymphoid cells that are susceptible to MMTV infection. These infected cells act as a reservoir for MMTV until the mice reach puberty, allowing mammary gland development. Actively dividing mammary epithelial cells then rapidly allow spread of the infection [37, 38, 39].

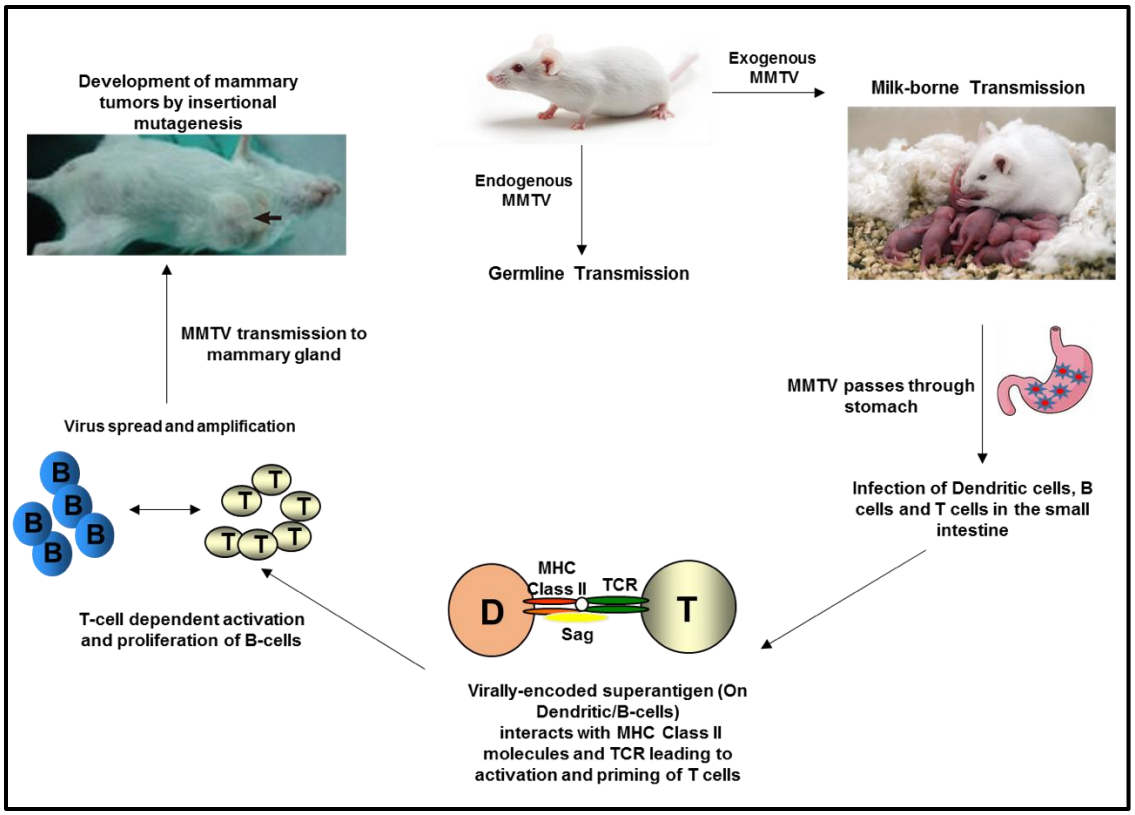

Figure 1.3: MMTV life cycle.

MMTV is transmitted by either the exogenous or endogenous routes. Exogenous MMTVs are transmitted through the milk to the small intestine, where DCs are infected and viral Sag is presented to T cells. Sag-mediated release of T-cell cytokines allows MMTV infection of B and T cells that carry virus to the developing mammary gland during puberty. Multiple cycles of viral replication in mammary epithelial cells lead to mammary tumor development [30].

The tissue specificity of MMTV expression is tightly controlled by regulation of virus transcription. Studies have shown that at least two transcription factors, Cux1 and SATB1, can bind the NRE region within the LTR and suppress transcription in non-mammary tissues [11, 40] (Fig. 1.4). SATB1 is absent in mammary tissues, whereas Cux1 undergoes cleavage during late pregnancy, thus relieving transcriptional repression of MMTV [41]. The hormone-responsive element (HRE) provides transcriptional responsiveness to glucocorticoids and also other steroid hormones, e.g. androgens and progesterone during pregnancy and lactation, thereby maximizing the virus production in milk [42, 43]. In addition, the mammary gland enhancer (MGE) in the LTR binds multiple transcription factors expressed in mammary epithelial cells to further upregulate hormone-induced transcription [44, 45].

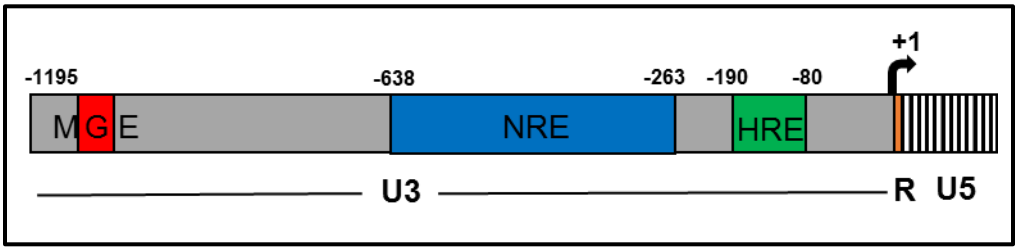

Figure 1.4: Transcriptional control elements in the MMTV LTR.

The transcription start site (+1) is located at the U3/R border and is denoted by an arrow. The mammary gland enhancer (MGE) is located at the 5' end of the LTR within the U3 region. The negative regulatory region (NRE) has binding sites for Cux1 and SATB1. The hormone response element (HRE) has binding sites for multiple steroid receptors including glucocorticoid receptor.

MMTV does not encode an oncogene, but instead causes mammary carcinomas by insertional mutagenesis near cellular proto-oncogenes. A number of common integration sites (CISs) have been reported in MMTV-induced tumors, particularly near *Wnt* and *Fgf* genes [46]. Because MMTV insertion into the cellular genome is relatively random, many integration events are required in mammary epithelial cells before insertion occurs near an appropriate proto-oncogene. The lack of targeted insertions and the subsequent selection of cells with MMTV-induced proto-oncogene activation results in a long tumor latency period of about six to nine months. Hence, mice that go through multiple rounds of pregnancy and lactation have higher levels of MMTV infection and an increased likelihood of developing mammary tumors [47].

1.1.4 Regulator of export/expression of MMTV mRNA (Rem)

Rem is a 301 amino-acid precursor protein that regulates the nuclear export of unspliced MMTV genomic RNA, as well as the post-export activities of all MMTV RNAs [23, 48]. Rem consists of a 98 amino-acid N-terminal signal peptide (SP) and a 203 amino-acid C-terminus (Fig. 1.5). Rem is translated from a doubly spliced mRNA that overlaps with the *env* gene. Since Rem and Env share the same open reading frame, both proteins encode for SP. SP contains many functional elements, such as a nuclear localization signal (NLS), a nucleolar localization signal (NoLS), an arginine-rich motif (ARM) and a nuclear export signal (NES). All of these motifs are shared by other retroviral proteins, such as HIV Rev, HTLV Rex, and human endogenous retrovirus type K (HERV-K) Rec [48].

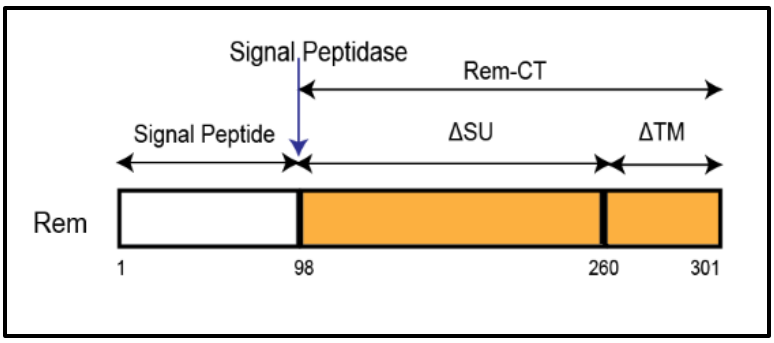

Figure 1.5: Diagram of the Rem precursor and its cleavage products.

Rem is synthesized as a precursor, which is cleaved by signal peptidase at the ER membrane into SP and Rem-CT. Rem-CT is a truncated version of the envelope surface and transmembrane domains.

SP generated from Env and Rem is longer than typical signal peptides and directs the precursor Rem to the endoplasmic reticulum (ER) membrane for translation. In the ER, signal peptidase cleaves the precursor protein into N-terminal SP and C-terminal protein (Rem-CT) [23, 50]. Uncleaved Rem is rapidly ubiquitylated, retrotranslocated to the cytoplasm, and subjected to ER-associated degradation (ERAD). In contrast, SP retrotranslocates to the cytoplasm, where it subverts ERAD to avoid proteasomal degradation [24]. Recent data from this laboratory indicates that SP is not ubiquitylated and binds to the p97 ATPase for membrane extraction into the cytosol [49]. Subsequently, SP traffics to the nucleus,where it binds unspliced MMTV mRNA for cytoplasmic export using the Crm1 pathway. In the cytoplasm, MMTV genomic RNA has two fates: translation into the unglycosylated viral proteins or incorporation into virus particles. Therefore, SP is a regulatory protein that is required for MMTV replication.

In contrast to SP, Rem-CT is a truncated version of Env and consists of the N-terminal part of surface (SU) protein and the C-terminal part of the transmembrane domain (TM) (Fig. 1.2). Rem-CT has two distinct populations. A fraction of Rem-CT utilizes ERAD to retrotranslocate to the cytoplasm, whereas another fraction of the protein leaves the ER and follows a unique trafficking pathway. Although it lacks a transmembrane domain, a portion of Rem-CT does not follow the typical secretory pathway, but may use endosomal trafficking (Wendy Xu, unpublished data). The functions of Rem-CT are yet unknown.

1.2 TYPE –B LEUKEMOGENIC VIRUS (TBLV)

TBLV is a naturally occurring variant of MMTV that causes T-cell lymphomas in mice, typically with shorter latencies than MMTV-induced mammary tumors [51-55]. The main differences between the TBLV and MMTV genomes lie within their LTRs (Fig. 1.6). The TBLV LTR has a deletion of 443 nucleotides in the negative regulatory element (NRE). Binding of Cux1 and SATB1 transcription factors to the NRE represses viral expression in a number of cell types, including lymphoid cells [11, 41]. In addition, TBLV has a substitution of 124 nucleotides in the U3 region of LTR, which results in triplication of a 62-bp element comprised of sequences flanking the NRE deletion [56-63]. This triplicated region serves as a T-cell specific enhancer element and provides binding sites for transcription factors, such as c-Myb, Runx1 and GR [64-65]. Two unidentified protein complexes, nuclear factor A (NF-A) and nuclear factor B (NF-B), and the transcriptional coactivator Aly are also critical for TBLV T-cell enhancer function in tissue culture

experiments. Furthermore, the proto-oncogenes *c-myc*, *tblvi1* and *Rorc* are three common integration sites identified in TBLV-induced lymphomas. Approximately 90% of TBLV-induced tumors overexpressed *c-myc* gene as a result of insertional activation by TBLV LTR upstream or downstream of the promoter in either orientation [66-67].

Deletion and substitution events within the TBLV LTR result in truncation of the *sag* open reading frame. This resulting truncated Sag protein retains the transmembrane region, part of the extracellular region, and the MHC-class II protein-binding motif. However, TBLV-encoded Sag lacks the C-terminal region that is required for its interaction with the TCR [68-69]. In contrast to MMTV-encoded Sag and its necessity for virus-induced mammary cancers, loss of TCR interaction is dispensable for induction of T-cell tumors in mice [70]. An early frameshift mutation within the TBLV truncated *sag* gene is compatible with the ability of the virus to cause T-cell lymphomas. Thus, TBLV is a Sag-independent variant of MMTV.

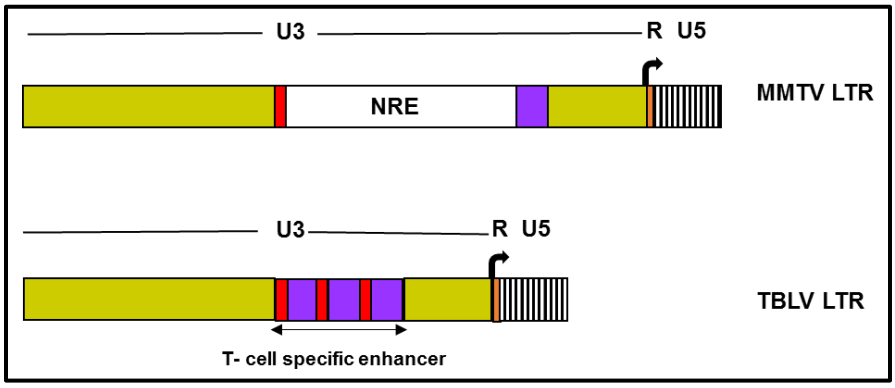

Figure 1.6: Comparison of MMTV and TBLV LTRs.

Both MMTV and TBLV LTRs consist of U3-R-U5 regions and have the transcription start site at the U3/R border. The 5' 18 bp and 3' 44 bp sequences flanking the NRE constitute the triplicated T-cell enhancer in the TBLV-LTR enhancer. These sequences are shown in red and purple, respectively.

1.3 APOBEC Family of Cytidine Deaminases

The APOBEC family of enzymes deaminate cytidine to uracil on single-stranded substrates by a zinc-mediated hydrolytic mechanism [71-83]. This family has eleven members in humans: APOBEC1 and AID (tightly linked on chromosome 12), APOBEC2 (chromosome 6), seven APOBEC3s (chromosome 22) and APOBEC4 (chromosome 1) (Fig. 1.7). AID and APOBEC2 have been identified in genomes of various bony fish and are considered the ancestral members, whereas the rest likely appeared as a result of gene duplication [84-88]. APOBEC1 (apolipoprotein B messenger RNA editing catalytic polypeptide 1) was the first member of this family to be identified that catalyzes cytidine-to-uridine editing of *APOB* mRNA in human intestine [89]. This modification generates a premature stop codon in the mRNA, giving rise to a truncated form of APOB protein called APOB48 that is essential for the transport of dietary fat [90]. APOBEC2 is expressed in

skeletal muscles and heart, whereas APOBEC4 is found in the testicles, but has no known deamination activity [91-92].

Activation-Induced Cytidine Deaminase (AID) is another member of this family that acts on the immunoglobulin (Ig) locus during transcription in germinal center (GC) B cells [93-94]. This enzyme has two major effects that result in antibody affinity maturation: hypermutation of the Ig variable regions and class-switch recombination (see 1.3.1). Both APOBEC1 and AID are catalytically active inside the nucleus but are localized predominantly in the cytoplasm to limit access to the nuclear compartment. AID activity is regulated at various transcriptional and post-translational levels. Off-target effects of AID have been implicated in B-cell lymphomas [95]. Overexpression of AID has also been associated with development of T-cell lymphomas due to increased point mutations within T cell receptor and *c-Myc* gene, in transgenic mice constitutively and ubiquitously expressing AID [96].

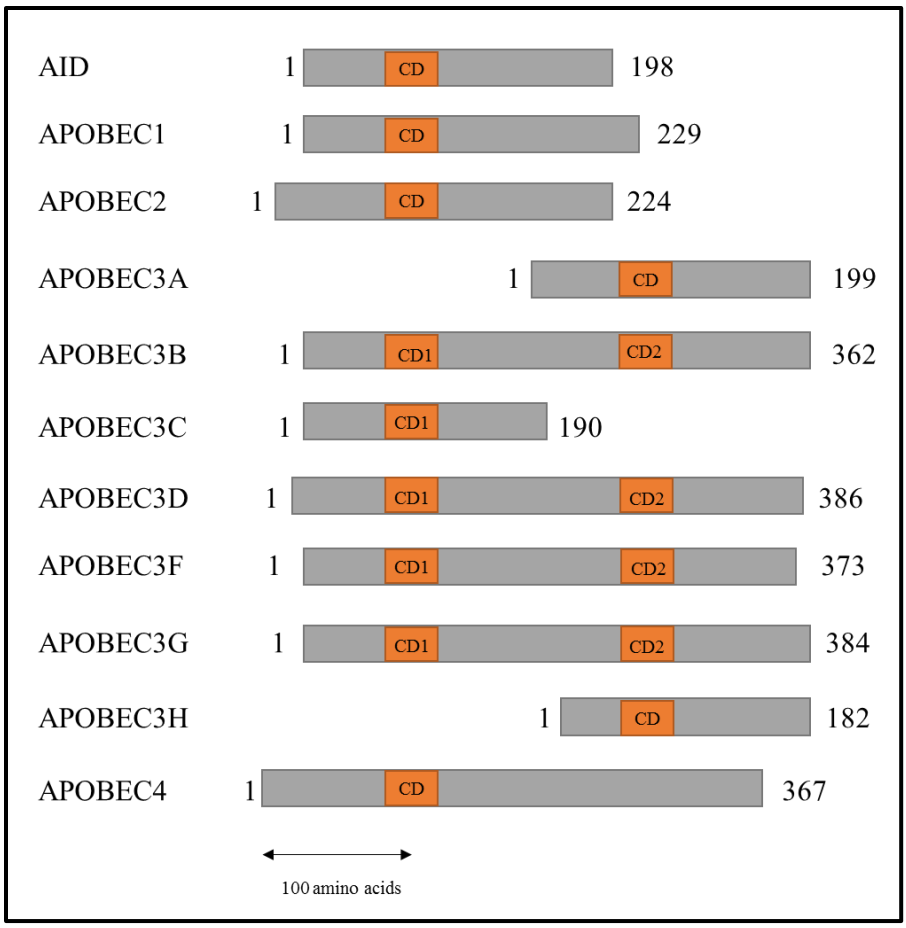

Figure 1.7: Schematic diagram of human APOBEC proteins.

APOBEC proteins are drawn to scale and total number of amino acids is shown to the right of each protein. Catalytic domains or the zinc binding domains (CD) are shown in orange and are present in single or double copies. For proteins with two CD copies, the N- and C-terminal domains are named CD1 and CD2, respectively [99].

APOBEC3 enzymes have been detected exclusively in mammals and restrict exogenous viruses as well as endogenous transposable elements (see 1.3.2). Humans have seven APOBEC3 (3A, 3B, 3C, 3D, 3F, 3G and 3H) genes that are expressed in a wide variety of tissues, whereas mice express only one copy of this gene [97]. APOBEC3 proteins vary in their subcellular localization: APOBEC3B (A3B) localizes to the nucleus;

APOBEC3D (A3D), APOBEC3F (A3F) and APOBEC3G (A3G) are present in the cytoplasm; whereas APOBEC3A (A3A), APOBEC3C (A3C) and APOBEC3H (A3H) are found in both - nucleus and cytoplasm. [84, 98].

1.3.1 Activation-Induced Cytidine Deaminase (AID) and virus restriction

Activation-Induced Cytidine Deaminase (AID) is a key regulator of B-cell diversity and plays a central role in development of a diverse repertoire of high affinity antibodies [100-102]. On encountering an antigen (Ag), B cells undergo division and diversification of antibody production in germinal centers (GC) of secondary lymphoid organs. AID catalyzes diversification of antibodies by somatic hypermutation (SHM) and class switch recombination (CSR) of Ig genes. This enzyme is known to have a preference for single-stranded DNA, which becomes accessible during transcription of the Ig loci [103]. Cytidine within the WRC or WRCY (W = A/T, R = purine, Y = C/T) motif context is preferred by AID for editing. During SHM, AID introduces point mutations in the antigen-binding variable region of an antibody. These mutations result in antibodies with increased antigen affinities, which allows B cells expressing edited B-cell receptors (BCRs) to proliferate and be subjected to positive selection during affinity maturation.

In contrast to hypermutation, which occurs on both Ig heavy and light chains, CSR occurs exclusively on the Ig heavy chain (IgH). AID-induced mutations in the Ig switch regions located upstream of each IgH constant region are processed as double-stranded breaks. Processing results in recombination events that replace the constant region of the mu heavy chain (Cμ) with one of the downstream IgH constant regions (Cγ, Cε,

Cα). This event controls antibody effector functions without affecting their antigen-binding affinity [104-110]. Deficiency in SHM and CSR leads to Hyper-IgM syndrome, which is associated with increased susceptibility to infections [111]. However, AID has the ability to target non-Ig genes as well, with cMyc-IgH translocations being prominent in B-cell malignancies. Thus, unregulated AID poses a threat to genomic integrity and hence is controlled at several levels including transcription, cellular localization and post-translational modifications [112-126]. At the level of mRNA stability, two miRNAs have been shown to regulate AID expression. The cellular distribution of AID is actively controlled by nuclear localization and nuclear export sequences. Various cofactors influencing nuclear access and cytoplasmic retention have also been identified. Post-translationally, proteasomal degradation and modifications involving phosphorylation affect the catalytic activity of AID.

More recent research supports roles for AID that are not limited to B-cell differentiation and antibody diversification. AID has been implicated in deamination of methylated cytidines leading to demethylation and epigenetic reprogramming [127]. Knockdown experiments showed that AID expression is required for *OCT4* and *NANOG* gene demethylation and initiation of nuclear reprogramming during pluripotency in human somatic cells. A recent study has also shown AID expression in a unique subset of CD4+ T cells, although with no known function [128].

AID has been shown to contribute to innate immunity by restricting the infectivity of certain viruses, particularly those with a DNA genome or intermediate. Hepatitis B Virus (HBV) is a DNA-containing virus that replicates in hepatocytes through

an RNA intermediate using reverse transcriptase [129]. AID has been reported to recruit RNA exosomes to the HBV ribonucleoprotein complex resulting in degradation of HBV RNA [130]. In addition, Abelson MuLV, which transforms immature B cells, has been shown to induce AID expression in non-germinal center B cells. Virus infection triggers the DNA-damage response pathway resulting in restricted proliferation [131]. However, neither of these cases result in AID-induced viral mutations. On the other hand, some B-cell tropic viruses such as Epstein-Barr Virus (EBV) and Kaposi's sarcoma herpes virus (KSHV) can inhibit AID, which is activated as a result of viral infection. EBV upregulates a host microRNA (miR-155) that targets AID [133] and KSHV genome encodes for two microRNAs that antagonize AID [134]. Apart from exogenous viruses, AID has also been implicated in restriction of endogenous long interspersed endogenous element (Line-1 or L1) transposons [132] by a deamination-independent mechanism. AID has the highest expression in differentiated B cells, which are indispensable for establishment of MMTV-infection [193]. Therefore, it is likely that MMTV has a mechanism for thwarting AID-mediated viral restriction during replication in B cells.

1.3.2 Human APOBEC3 and virus restriction

Human A3D (hA3D), hA3F, hA3G and hA3H have been shown to restrict Vif-deficient HIV-1(Fig. 1.8). In the absence of Vif, hA3 proteins form a cytoplasmic complex with HIV-1 Gag and get incorporated in the capsid of the budding viral particle [135-140]. After target cell infection, the viral capsid uncoats and RNA genome undergoes reverse transcription in the cytosol. Single-stranded intermediates on the viral minus-strand DNA

serve as a substrate for cytidine deamination by hA3 proteins that have been encapsidated in the viral particle. Uracil-containing minus strand complementary DNA (cDNA) functions as a template for second-strand synthesis and results in G to A mutations on the plus strand. Mutations lead to proviral instability or result in the introduction of stop codons, which affect the ability to produce a functional virus after integration into the host genome [141-145]. Studies have also shown that deamination-defective hA3G and hA3F can also restrict HIV-1 by binding genomic RNA and blocking access to reverse transcriptase [146-150]. However, HIV-1 Vif can counteract hA3 proteins by recruiting CBF-β to form a stable E3 ubiquitin ligase complex. This complex triggers polyubiquitylation and subsequent degradation of hA3 proteins, thus preventing virion incorporation in the producer cell [151-153]. All mammals encode for at least one A3 and other lentiviruses such as simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV) and bovine immunodeficiency virus (BIV) encode Vif proteins that degrade A3 enzymes of their respective hosts [154-156]. Therefore, viral antagonists of cytidine deaminases are known to be host-specific.

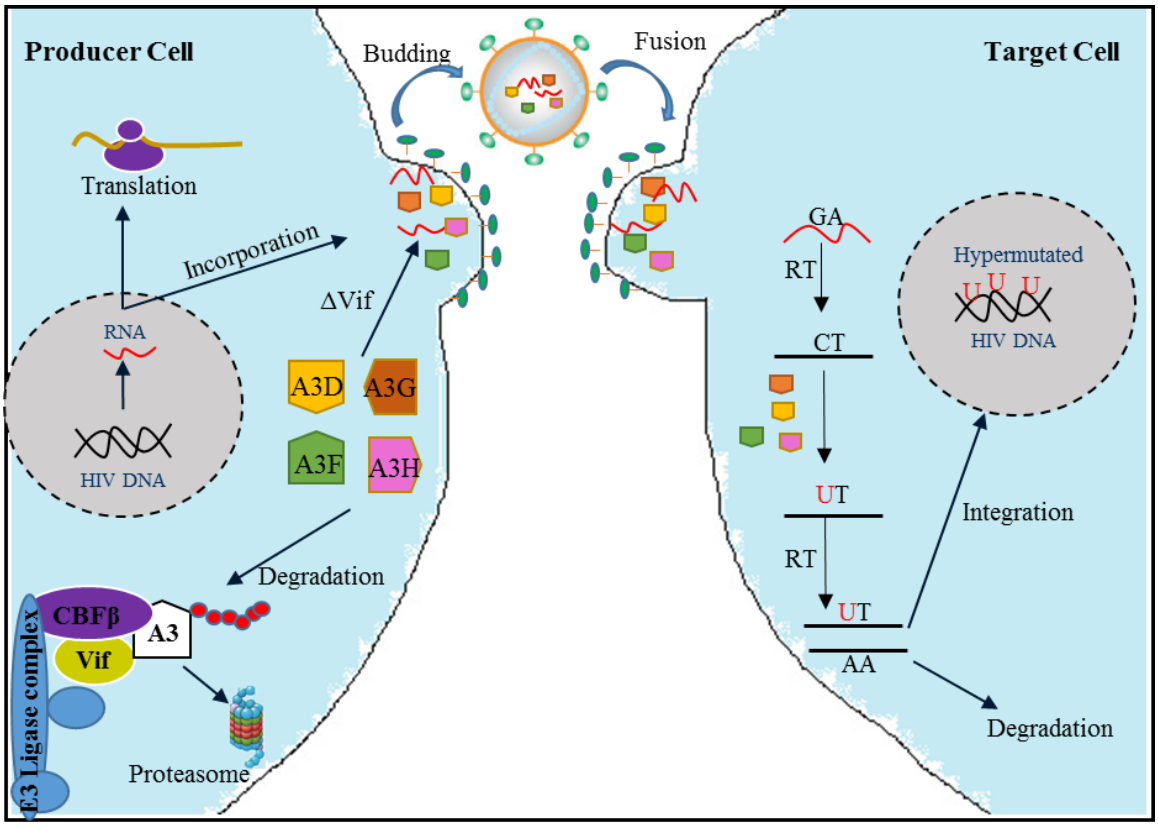

Figure 1.8: Model for HIV-1 restriction by human APOBEC3 proteins and Vif-mediated antagonism.

In the virus producing cells, Vif acts as an adaptor molecule between the E3 ligase complex and specific hA3 proteins, leading to polyubiquitylation (shown by red circles) of the deaminases and their subsequent degradation by the 26S proteasome. In the absence of Vif, different hA3 proteins get incorporated in the budding viral particle along with viral RNA. After fusion of the viral particle with a target cell and initiation of reverse transcription (RT), hA3 proteins deaminate cytidines to uracils on single-stranded viral minus strand cDNA. This uracil on the template strand results in incorporation of adenine on the complementary plus strand, giving rise to guanine to adenine mutations. These hypermutated cDNAs are either targeted for degradation or integrated into the host genome. Many integrated proviruses have multiple stop codons that prevent viral protein function, which further contribute to lower viral titers [97].

Human T-cell leukemia virus type-1 (HTLV-1) has been shown to be restricted by hA3 enzymes via both deaminase-dependent and independent mechanisms. However, HTLV-1 is considerably more resistant than HIV-1 to hA3D, hA3F and hA3G, although HTLV-1 also is susceptible to hA3A and hA3B [157-159]. HTLV-1 does not encode a Vif accessory protein that targets deaminases for degradation [160-161]. Instead, a unique C-terminal extension of HTLV-1-encoded nucleocapsid (NC) protein has been demonstrated to antagonize A3 enzymes by preventing their encapsidation in the viral particle. In addition to exogenous retroviruses, A3 proteins inhibit mobility of LTR-containing retroviruses (HERVs) as well as non-LTR retrotransposons (LINEs and SINEs) in humans by sequestration in the cytoplasm [162-164]. Apart from retroelements, studies have also documented hA3-induced mutations in DNA viruses, e.g. Hepatitis B virus (HBV) and Human papilloma virus (HPV) [165-166].

1.3.3 Murine Apobec3 and virus restriction

Several studies have described the effects of murine Apobec3 (mA3) on different strains of the gammaretrovirus Murine Leukemia Virus (MuLV). Mice lacking mA3 have 10-100 fold higher levels of Friend (F)-MuLV infection in both spleen and bone marrow [167], suggesting that this enzyme is active in lymphoid cells. In addition, the mA3 locus has also been shown to encode for Recovery from Friend Virus 3 (*Rfv3*) gene, which has been associated with regulating the neutralizing antibody response to Friend virus (FV) infection [168]. FV is a complex of the defective spleen focus forming virus (SFFV) and the non-defective F-MuLV [177-178]. Studies of another gammaretrovirus that infects

lymphoid cells, Moloney (M)-MuLV, revealed that infection of mA3-knockout mice accelerated T-cell lymphoma induction. Establishment of infection was significantly delayed in spleens and bone marrow-derived cells compared to the same tissues of mA3-expressing C57BL/6 mice [169]. Moreover, infectivity of M-MuLV with packaged mA3 was reduced further by mA3 expressed in the target cells that were being infected [170]. Interestingly, in experiments carried out in AKR strain of mice that express an endogenous MuLV (AKV), mA3-mediated restriction was accompanied by G to A hypermutation of the viral genome [171]. The mA3 enzyme deaminates cytidines in the TYC sequence motif (Y = T/C). Significant levels of cytidine deamination by mA3 have not been reported for any of the exogenous MuLVs [172-173].

Although mA3 caused G to A mutations in HIV in co-expression experiments in tissue culture, this enzyme is resistant to Vif-mediated degradation [174]. Multiple studies have shown that MuLVs are more resistant to mA3 than to hA3 enzymes. Primary protection of MuLVs from mA3 inhibition is contributed by a glycosylated form of the capsid precursor protein Gag (Glyco-Gag). This longer form of Gag has an additional 88 amino acids at the N-terminus and has been reported to confer stability to the virus capsid. Glyco-Gag-mutant viruses cause the capsid to be loosely formed and more accessible to mA3-mediated blocks to reverse transcription [170,175]. A second layer of protection is provided by MuLV protease, which cleaves the mA3 allele encoded by BALB/cJ mice within exon 5. C57BL/6 mice encode mA3 (mA3Δexon5) that is resistant to this cleavage [176]. These results suggest that there are strain-specific differences in mA3-mediated MuLV restriction.

1.4 MMTV RESTRICTION BY MURINE APOBEC3

The first evidence of mA3-mediated viral restriction was demonstrated *in vivo* by studying MMTV [179]. Murine Apobec3 (mA3) mRNA is detectable in lymphoid cells, macrophages and dendritic cells. However, the highest levels of mA3 have been reported in lymphoid cells, which are known to be indispensable for establishment of MMTV infection *in vivo*. Mice with targeted deletion of the murine *Apobec*3 gene on C57BL/6 background have higher levels of MMTV DNA and are more susceptible to infection by the RIII strain of MMTV [179]. Very low levels of mA3-induced G to A mutations were detected in these MMTV sequences, suggesting a deamination-independent means of restriction [180-181].

BALB/cJ mice encode a longer mA3 isoform consisting of 9 exons, whereas C57BL/6 mice lack exon 5 and encode a ~3.6 kDa smaller protein [182-184]. Additionally, BALB/cJ mice express lower levels of mA3 compared to C57BL/6 mice and hence may be less restrictive to murine viruses. Virions isolated from mammary tissues of MMTV-infected mice showed that both allelic variants of mA3 are incorporated into the particle and retain their deamination activity. Subsequently *in vitro* infection assays revealed that the major means of mA3-mediated restriction of MMTV was by inhibition of reverse transcription and not cytidine deamination [180].

Further, mA3 expression was observed in mammary epithelial cells (MECs), where it can get incorporated in the virus particle to restrict milk-borne transmission. Experiments conducted to test the infectivity of virus isolated from milk revealed that virus from mA3-knockout mice infected cells at higher levels than that isolated from wild-type mice [182].

Hence, mA3 potentially restricts MMTV at various steps during virus replication and transmission *in vivo*.

1.5 RATIONALE FOR THIS STUDY

MMTV is a betaretrovirus that causes mammary carcinomas in mice and has been successfully used as a model system to study human breast cancer. Unlike many other murine viruses, MMTV has a complex genome organization similar to human retroviruses such as HIV and HTLV. Apart from structural proteins that form the virus particle, MMTV encodes additional accessory and regulatory proteins that facilitate its pathogenesis.

Prior to causing mammary gland tumors by insertional mutagenesis, MMTV establishes its infection in the lymphoid population. Dendritic cells, B cells and T cells serve as a reservoir for the virus until mice reach puberty, and the virus has access to actively dividing mammary epithelial cells [37]. MMTV has evolved ways to both evade and use the host immune system. Thus, studying the viral pathogenesis will help us identify mechanisms of innate and adaptive immune evasion.

MMTV encodes a 33 kDa regulatory protein, Rem. Signal peptide (SP) generated from the N-terminus of Rem functions analogous to HIV- Rev protein. SP binds to a cis-acting element within the viral RNA and aids in exporting it to the cytoplasm. SP is also produced from MMTV envelope protein. Hence, this raises the question of how expression of full - length Rem can provide an advantage to the virus *in vivo*.

Previous studies have shown that MMTV replicates more efficiently in the absence of the restriction factor murine apobec3 (mA3). However, viral sequences analyzed from

these studies lacked detectable levels of mA3-induced cytidine deamination. We were motivated to infect BALB/cJ and C57BL/6 mice with virus that lacks Rem expression in order to ascertain if Rem affects the innate responses to MMTV infection.

This study demonstrates that absence of Rem makes the MMTV genome more susceptible to mutations occurring in motifs known to be recognized by the Apobec family members, namely AID and mA3. An independent study performed with the thymotropic variant TBLV also showed the same mutational phenotype in BALB/cJ mice, suggesting that Rem is an antagonist of AID and mA3 possibly in different cell types. Our results thus point towards a Vif-like function of Rem in protecting the viral genome from the activity of cytidine deaminases in the lymphoid population. Although AID has been shown to restrict retroviruses, these experiments provide the first example of a retroviral genome mutated by AID. Surprisingly, AID-knockout (AID-KO) mice demonstrated reduction in motifs known to be recognized by mA3, implicating AID as a regulator of other cytidine deaminases. Since AID overexpression has been associated with multiple human cancers, elucidation of the mechanism by which Rem antagonizes AID will be useful for treatment of human diseases.

Chapter 2: Materials and Methods

2.1 CELL LINES

XC rat fibroblast cells were grown in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Carlsbad, CA) supplemented with 5% heat-inactivated (30 min at 56°C) fetal bovine serum (FBS) (Sigma), 100 Units/ml penicillin, 100 μg/ml streptomycin (Invitrogen), 2 mM L-glutamine and 50 μg/ml gentamycin sulfate (Elkins-Sinn, Inc., Cherry Hill, NJ). Derivation of stable XC/MMTV transfectants has been described previously [185-186]. Jurkat human T-lymphoma cells were maintained in suspension using Roswell Park Memorial Institute (RPMI) 1640 medium (Invitrogen) containing 5% FBS, 100 Units/ml penicillin, 100 μg/ml streptomycin, 2 mM L-glutamine and 50 μg/ml gentamycin sulfate. Human embryonic kidney cells (293T) were maintained in DMEM containing 10% FBS, 100 Units/ml penicillin, 100 μg/ml streptomycin, 2 mM L-glutamine and 50 μg/ml gentamycin sulfate. All cell lines were maintained at 37°C in a humidified atmosphere containing 7.5% CO_2.

2.2 MICE

Weanling (4-6 weeks old) mice were injected intraperitoneally with 2 X 10^7 XC cells producing MMTV-WT or MMTV-SD or 2 X 10^7 Jurkat T cells producing TBLV-WT or TBLV-SD [70,185]. Prior to inoculation into mice, the cells were washed twice with phosphate-buffered saline (PBS) and resuspended in 500 μl of PBS. Aliquots of the cells inoculated into mice were analyzed by Western blotting to verify the expression of viral

proteins. XC cells stably transfected with MMTV were induced for 24 h before injection into mice by supplementing the culture medium with 10^{-6} M dexamethasone (Sigma Chemical Laboratories, St. Louis, MO). Jurkat cells were not treated with Dex since they lack functional glucocorticoid receptors. AID-KO mice on the BALB/cJ background were derived by 10 backcross generations with AID-KO mice on the C57BL/6 background obtained from Dr. T. Honjo, Kyoto University [93] (kindly provided by Dr. M. Nussenzweig, Rockefeller University). A total of 80 markers were used to distinguish between BALB/cJ and C57BL/6 backgrounds. All mice were confirmed to have the correct genotype by PCR typing of tail DNA as described [32]. Injected mice were monitored for the appearance of T-cell lymphomas or mammary carcinomas at least weekly, and the resulting tumors were harvested and stored frozen at -70^{o}C. AID-GFP transgenic mice on the C57BL/6 background were obtained from Jackson Laboratories [218]. Mouse breeding, maintenance, inoculations and PCR typing were performed by Mary Lozano. All animal experiments were approved by the UT Austin IACUC.

2.3 GENOMIC DNA PREPARATION

Tissue fragments obtained by crushing frozen tissue were incubated in lysis buffer [75 mM NaCl, 20 mM Tris-HCl, pH 7.4, 25 mM EDTA, pH 8.0, 0.5% (w/v) SDS, 100 μg/ml fresh proteinase K] at 37°C for 1 to 2 days. If the lysates were very viscous, additional proteinase K and lysis buffer were added during this incubation. The lysates were then extracted with phenol-chloroform, and the aqueous phase was collected in a separate tube. DNA was adjusted to 0.2 M NaCl and precipitated using 70% ethanol. The genomic DNA

was spooled out into an Eppendorf tube using a Pasteur pipette, and the majority of ethanol was removed. The DNA was re-suspended in a solution of 10 mM Tris-HCl, pH 7.4 and 0.1 mM EDTA (TE 10:0.1) and stored at 4°C.

2.4 PLASMID PREPARATION

Large-scale plasmid preparations were performed by alkaline lysis and cesium chloride (CsCl) gradient centrifugation. Briefly, 500 ml of Luria-Bertani (LB) broth [1% tryptone w/v, 0.5% yeast extract (w/v), 0.5% NaCl w/v, pH 7.4] containing appropriate antibiotic was inoculated with 25 ml of an overnight culture of bacterial transformant growth in LB. The culture was grown in a two liter Erlenmeyer flask and grown at 37°C with shaking. Bacteria were pelleted by centrifugation (Avanti J-E centrifuge, Beckman) at 7,000 rpm in a JA10 rotor (Beckman) for 10 min at 4°C. The pellet was resuspended in 7 ml of Solution I (50 mM glucose, 25 mM Tris-HCl, pH 8.0, and 10 mM EDTA) and transferred to a 50 ml conical tube. Lysozyme [stock concentration of 10 mg/ml (Sigma)] was added to a final concentration of 1 mg/ml, and the suspension was incubated at room temperature (RT) for 10 min. Sixteen ml of freshly prepared Solution II [0.2 N NaOH and 1% sodium dodecyl sulfate (SDS)] was added to lyse the bacterial cells by gently inverting the tube 3-4 times, followed by a 10 min incubation at RT. The alkaline reaction was neutralized by adding 12 ml of cold Solution III (3 M potassium acetate, pH 4.8, in glacial acetic acid) and mixed thoroughly by shaking the tube several times. The mixture was incubated on ice for 10 min, followed by centrifugation in an Allegra X-15R centrifuge (Beckman Coulter) at 3,800 rpm for 30 min at 4°C. Chromosomal DNA was pelleted and

the supernatant was filtered through cheesecloth into a new 50 ml conical tube. Isopropanol was added to the tube and incubated at RT for 15 min. DNA was precipitated by centrifugation at 3,800 rpm for 30 min at RT in an Allegra X-15R centrifuge. The DNA pellets were washed with 70% ethanol and air dried for 10 min. Pellets were resuspended in 3.5 ml of 10 mM Tris-HCl, pH 7.4, and 10 mM EDTA (TE 10:10). CsCl gradient ultracentrifugation was then used to purify the DNA by removing RNA and proteins.

CsCl (4.7 g) and ethidium bromide (EtBr) (0.78 ml of a 5 mg/ml solution) were added to DNA dissolved in TE 10:10. Insoluble debris (RNA and protein) was removed by centrifugation at 3,500 rpm for 15 min at 4°C in an Allegra X-15R centrifuge. The supernatant was transferred to a Beckman Optiseal tube, and the suspension was separated by centrifugation in a Beckman L7-65 ultracentrifuge at 50,000 rpm for 18 h at 20°C using an NVT 65.2 rotor. The plasmid band was transferred to a 15 ml conical tube with a help of an 18 gauge needle and EtBr was extracted 3-4 times until a clear solution was obtained. The extraction was done by using 2 volumes G-50 buffer (20 mM Tris-HCl, pH 7.4, 0.1 M NaCl, and 2 mM EDTA) saturated n-butanol. Subsequently, DNA was precipitated by addition of ice-cold 100% ethanol, 0.2M NaCl and centrifugation at 3,800 rpm for 30 min at 4°C in an Allegra X-15R centrifuge. Following a 70% ethanol wash and air-drying, the DNA pellet was resuspended in 2 ml of TE (10:10). RNA contaminants were removed by incubation at 37°C for 30 min with 100 µg/ml of pancreatic RNase A (Sigma) and 200 U of RNase T1 (Sigma). The solution was incubated with 100 µg/ml of proteinase K (Sigma) in the presence of 0.5% SDS for 30 min at 37°C. Following phenol:chloroform extraction, the solution was dialyzed using a 10,000 molecular-weight cut-off (MWCO) Slide-A-

Lyzer® dialysis cassette (Pierce, Rockford, IL) against 10 mM Tris-HCl, pH 7.4 and 1 mM EDTA (TE 10:1) with three buffer changes. Plasmid DNA was recovered by ethanol precipitation and resuspended in TE 10:0.1 following a 70% ethanol wash.

Table 2.1: Summary of plasmids used in this study.

Plasmid	Details
pHYB MMTV-WT	Infectious MMTV-expressing and dexamethasone-inducible molecular clone [186].
pHYB MMTV-SD	pHYB-MMTV-WT molecular clone carrying a mutation in the second Splice Donor (SD) site [185].
pHYB TBLV-WT	Infectious TBLV-expressing molecular clone [70].
pHYB TBLV-SD	pHYB-TBLV-WT molecular clone carrying a mutation in the second Splice Donor (SD) site [70].
GFP-Rem	N-terminally EGFP-tagged Rem expression plasmid.
Rem	Untagged Rem expression plasmid made by introducing Rem coding region into EGFPN3 vector lacking GFP.
pGL3	Firefly luciferase reporter vector from Promega
pHM*RLuc*	Rem-responsive *Renilla* luciferase reporter vector [23]
pcDNA3.1	Control vector from Addgene used in transfection assays.
CMV MMTV - WT	CMV promoter-driven infectious MMTV-expressing molecular clone obtained from Dr. Farah Mustafa [189].
CMV MMTV - SD	CMV MMTV-WT molecular clone carrying a mutation in the second Splice Donor (SD) site.
mA3-HA	HA-tagged C57BL/6 Apobec3 cloned into pcDNA3.1 obtained from Dr. Nathaniel Landau [174].
mAID-GFP	C-terminally GFP-tagged murine AID expression plasmid obtained from Dr. Frederick Alt [187].
pcDNA3.1-mAID	Untagged murine AID expression plasmid cloned from pMX-AID.
pMX-AID	Untagged AID expressed in a retroviral vector provided by Dr. Kevin McBride [188].
pHITG	CMV promoter-driven VSV-G expression plasmid provided by Dr. Glen Gaulton.
pHIT60	CMV promoter-driven Moloney MuLV Gag-Pol expression construct provided by Dr. Glen Gaulton.

2.5 PCR

DNA extracted from tumors induced by MMTV-WT and MMTV-SD were used for PCR with primers *env*7254(+) (5'-ATCGCCTTTAAGAAGGACGCCTTC T-3') and C3H LTR420(-) (5'-GATTCATTTCTTAACATAGTAAC-3') for the envelope region. For DNA extracted from tumors induced by TBLV-WT and TBLV-SD, LTR9604(-) (5'-GGA AACCACTTGTCTCACATC-3') was used as the reverse primer for *env* gene amplification. PCR parameters were: 94ºC for 1 min, then 10 cycles at 94ºC for 10 sec, 53ºC for 30 sec, and 68ºC for 2 min followed by 25 cycles of 95ºC for 15 sec, 50ºC for 30 sec, and 68ºC for 2 min, and a final incubation at 68ºC for 7 min. The *c-Myc* gene was amplified using primers *c-Myc*(+) (5'-ATGCCCCTCAACGTGAACTT-3') and *c-Myc*(-) (5'-AGGAGGTCCATCCAACCTCT-3'). PCR parameters were: 94ºC for 1 min, then 35 cycles at 94ºC for 30 sec, 58ºC for 30 sec, and 68ºC for 2 min, and a final incubation at 68ºC for 5 min. All reactions were performed with JumpStart RED Accutaq LA polymerase (Sigma-Aldrich) in the supplied buffer, 500 ng of tumor DNA, 50 pmol of each primer, and 0.5 mM deoxynucleoside triphosphates in 20 µl.

Semi-quantitative PCR was performed with 50ng DNA. *Mtvr2*(+) (5'-TCTGGG ATCCGCTTCCTCAT-3') and *Mtvr2*(-) (5'-CCAGTCCTTGGCCCTCATTTA-3') or *pol*4235(+) (5'-GAAGAGAGCAATAGCCCTTG-3') and *pol*5835(-) (5'–GATGATGTAGTGCGTGGC-3') primers were used for PCR. Levels of the *pol* gene product were normalized with levels of *Mtvr2*. DNA from MMTV-infected BALB/cJ mice was used as a control.

Proviruses from MMTV and TBLV-induced tumors amplified with virus specific primers were cloned using pGEM–T Easy Vector System 1 (Promega). Briefly, PCR products were separated on 1% agarose gel and the appropriate band was purified with DNA Clean and Concentrator Kit (Zymo Research). Purified DNA amplicons were ligated to pGEM-T Easy Vector with T4 DNA Ligase at 16^{o}C overnight. The ligation mixture was transformed into *E.coli* (DH5alpha) and DNA was extracted from resultant colonies using GenElute Plasmid Miniprep Kit (Sigma). DNA was sequenced with *env* (7254+) and *env* (8506-) (5’ – GCACTTGGTCAAGGCTCTTCG – 3’) primers.

Table 2.2: Summary of primers used in this study.

Primer	Sequence
*env*7254(+)	5’ATCGCCTTTAAGAAGGACGCCTTCT 3’
C3H LTR420-	5’GATTCATTTCTTAACATAGTAAC 3’
LTR9604(-)	5’GGAAACCACTTGTCTCACATC 3’
c-Myc(+)	5’ATGCCCCTCAACGTGAACTT 3’
c-Myc(-)	5’AGGAGGTCCATCCAACCTCT 3’
Mtvr2(+)	5’TCTGGGATCCGCTTCCTCAT 3’
Mtvr2(-)	5’CCAGTCCTTGGCCCTCATTTA 3’
env(8506-)	5’ GCACTTGGTCAAGGCTCTTCG 3’
*pol*4235(+)	5’GAAGAGAGCAATAGCCCTTG 3’
*pol*5835(-)	5’GATGATGTAGTGCGTGGC 3’

2.6 RNA EXTRACTION AND RT-PCR

Total RNA was extracted from tissue culture cells using the TRI Reagent (Molecular Research Center, Inc). Culture medium was removed, and one ml of TRI Reagent was added directly to the culture dish to lyse the cells. Lysate was transferred to an Eppendorf tube, pipetted several times to homogenize and incubated for 5 minutes to allow complete dissociation of the nucleoprotein complexes. Chloroform (0.2 ml) per 1 ml of TRI Reagent was added, followed by a 2–3 min incubation. Samples were then centrifuged for 15 min at 12,000 × g at 4°C. The mixture separates into a lower red phenol-chloroform phase, an interphase, and a colorless upper aqueous phase. The aqueous phase containing the RNA was transferred to a new tube. Isopropanol (0.5 ml) per 1 ml of TRI Reagent was added to the aqueous phase and incubated for 10 min. RNA was concentrated by centrifugation at 12,000 × g at 4°C for 30 min. After washing with 75% ethanol, the pellet was resuspended in 100 to 200 µl of diethyl pyrocarbonate-treated water (0.1% DEPC treatment followed by autoclaving).

Purified RNA (10 µg) was treated with 3U of amplification-grade DNase I (Invitrogen) and 0.5 U of RNaseOUT RNase inhibitor (Invitrogen) for 4 h at 37ºC. The DNase I reaction was stopped by the addition of 5 µl of 25 mM EDTA and incubation at 72ºC for 15 min. Reverse transcription reactions contained oligo primer poly(dT_{17}) and deoxynucleotide triphosphates (dNTPs) at final concentrations of 2.5 pmol/µl and 1 mM, respectively. The mixture was then boiled for 5 min, followed by immediate incubation on ice for 5 min. The RNA was reverse transcribed with 400 U of Moloney murine leukemia virus (M-MLV) reverse transcriptase (Invitrogen), 5 U of RNaseOUT, and 10 mM

dithiothreitol in a 50 µl reaction mixture for 1 h at 37ºC. The complementary DNA was used with primers C3H 230(+) (5'-GTG AAT TCC ATC ACA AGA GCG GAA CGG AC-3') and C3H LTR420(-) (5'-GAT TCA TTT CTT AAC ATA GTA AC-3') in reactions with JumpStart REDTaq Reaction Mix (Sigma-Aldrich). The PCR parameters were: 94°C for 2 min and then 30 cycles of 94°C for 20 sec, 55°C for 30 sec, and 68°C for 6 min and an 8 min extension at 68°C. Reactions were analyzed on 1% agarose gels and stained with EtBr prior to visualization using an AlphaImager (Alpha Innotech).

2.7 TRANSFECTIONS

Calcium Phosphate Precipitation: HEK 293 and 293T cells were transfected by the calcium phosphate method. One day (24 h) before transfection 5 x 10^5 cells were seeded in 6-well tissue culture plates in 2 ml of media. On the day of transfection, total 6 µg DNA was mixed with 10 µl of $CaCl_2$ (final concentration of 0.25 M) and the volume was adjusted to 100 µl with water. This solution was added dropwise to a second tube containing 100 µl of 2X-HBS, pH 7 (280 mM NaCl, 10 mM KCl, 1.5 mM $Na_2HPO_4 \cdot 2H_2O$, 12 mM dextrose, 50 mM HEPES) while vortexing. The mixture was incubated at RT for 15 – 20 min and added to the cells dropwise. At 6 to 8 h post-transfection media was changed to remove the calcium phosphate precipitate. Cells were harvested after 48 h.

Lipofectamine: Lipofectamine 3000 (Invitrogen) was also used for transfection of 293T cells. The day before transfection 10^6 cells were seeded in 6-well plates. For each transfection, 3 µg of DNA was mixed with 2 µl/µg of P3000 Reagent, and DMEM without additives was added for a final volume of 125 µl. In a separate tube, 5 µl of Lipofectamine

3000 was added to DMEM without additives for a final volume of 125 µl. The diluted DNA was then added to diluted Lipofectamine at 1:1 ratio and incubated for 20 min at RT. The DNA-Lipofectamine 3000 Reagent complexes were added to cells dropwise. Cells were harvested after 48 - 72 h.

Electroporation: Jurkat T cells and XC rat fibroblast cells were transfected by electroporation. On the day of transfection, 1 x 10^7 cells were resuspended in 400 µl of serum-free RPMI. DNA (20-40 µg) was added to this mix and transferred to a 4-mm gap cuvette. A Gene Pulser II (Biorad) was used to electroporate Jurkat and XC cells at 260V, 950 µF and 260V, 750 µF, respectively. Cells were replated in 60 mm plates with 5 ml of complete media after 10-15 minutes of incubation at RT. MMTV expression in XC cells was induced by addition of 10^{-6}M Dex for 24 h. Cells were harvested at 48 h.

2.8 PREPARATION OF RETROVIRUS-CONTAINING CULTURE SUPERNATANTS

Culture supernatants were aspirated from 293T cells transfected in two 6-well plates at 48 and 72 h post-transfection. After a short centrifugation at 4,000Xg to remove any cell contaminants, the supernatant was passed through a 0.45 micron filter and incubated overnight with one-third volume of RetroX concentrator (Takara) at 4°C. Virus was pelleted by centrifugation at 1,500Xg for 45 minutes at 4°C and resuspended in PBS. For protease treatment, half of the PBS resuspension was incubated with subtilisin (Roche) at a final concentration of 10 µg/ml for 1 h at 37°C [190-191]. Subsequently, the samples were analyzed by Western blotting.

2.9 RETROVIRAL TRANSDUCTION

Pseudotyped mAID-expressing virus was produced in HEK293T cells by co-transfecting retroviral plasmids pMX-AID along with pHITG (expressing VSV glycoprotein G) and pHIT60 (expressing MuLV Gag-Pol) (see Table 2.1). Two days post-transfection, culture supernatants were centrifuged thrice at 524Xg for 5 min in a Beckman Allegra X-15R centrifuge to remove cell debris. Pseudotyped virus was concentrated using a Beckman SW55Ti rotor and centrifugation at 100,000Xg for 1 h at 4^0C. Supernatants were decanted, and viral pellets were resuspended in PBS. Concentrated virus was transduced into Jurkat cells by spinoculation at 800Xg for 1 h at RT in the presence of 8 μg/ml polybrene to obtain an MOI >1. Transduced cells were grown in the presence of 2 μg/ml puromycin for 5-6 days.

2.10 REPORTER GENE ASSAYS

Firefly and *Renilla* luciferase luminescence was measured with the Dual-Luciferase® Reporter Assay System (Promega) according to the manufacturer's instructions. Transfected cells were washed with 1X PBS and lysed in 1X Passive Lysis Buffer (Promega) by three cycles of freezing at -80ºC for 20 min and thawing in a 37ºC water bath. The lysates were clarified by centrifugation at 4,000 rpm in a Beckman Coulter Microfuge 18 Centrifuge for 10 min at RT. Bio-Rad Protein assay system was used to quantitate protein concentration. A total of 20 - 40 μg of lysate was used to obtain readings for both firefly and *Renilla* luciferase using a Turner TD-20e luminometer (Turner Designs, Inc., Sunnyvale, CA) that was set to 0 delay and 10 sec integration. Luciferase Assay Reagent II (LAR II) (Promega) (50 μl) was added to the sample, and the firefly luciferase

activity was determined. The *Renilla* luciferase activity was determined after the addition of 50 µl Stop and Glo Reagent (Promega). The firefly luciferase was expressed in relative light units (RLUs), and the relative values were normalized to 100 µg of protein.

2.11 *EX VIVO* INDUCTION OF AID

Spleens were removed aseptically and single cell suspensions were made in 5 ml of wash solution containing 1X PBS and 1% FBS. Cells were subjected to centrifugation at 1,200 rpm for 7 min. Red blood cells were then lysed in 3 ml of Red Blood Lysis Solution (9 parts NH_4Cl and 1 part Tris-HCl pH 7.65) for 5 min at RT. Wash solution (20 ml) was added and cells were pelleted again at 1,200 rpm for 7 min. The resulting pellets were re-suspended in 2 ml of wash solution and filtered through a 40 micron strainer into 35mm Petri-dishes. Cells were then counted, plated at 5 x 10^5/ml and grown in stimulation media (RPMI 1640 with L-glutamine from Sigma, 1% Antibiotic/anti-mycotic solution from Gibco, 1 mM Sodium Pyruvate from Sigma, 10% FBS from Biowest, 50 uM 2-mercaptoethanol from Invitrogen, 5 ng/ml IL-4 from Sigma and 20 ug/ml LPS from Sigma. Maximum AID expression was observed at day 4.

2.12 WESTERN BLOTTING

Whole cell extracts were prepared by adding either one volume of 2X-SDS-loading buffer [250 mM Tris-HCl, pH 6.8, 20% glycerol, 2% sodium dodecyl sulfate (SDS), 5% β-mercaptoethanol, 0.2% bromophenol blue] to cells in one volume of phosphate-buffered saline (PBS) or one volume of 6X-SDS loading buffer (0.35 M Tris-HCl, pH 6.8, 10%

SDS, 36% glycerol, 0.6 M DTT, 0.012% bromophenol blue) to cells in five volumes of PBS. Samples were boiled for 5 min for protein denaturation.

Total denatured protein (10 to 100 μg) from the extracts was resolved in 10 or 12% polyacrylamide gels containing 1% SDS and transferred to Optitron 0.45 μm nitrocellulose membranes in Western transfer buffer [39 mM glycine, 48 mM Tris base, 1% SDS, 20% (v/v) methanol] overnight at 120 mA at 4ºC. The membrane was incubated with 5% nonfat dry milk in Tris-Buffered Saline Tween-20 (TBS-T) buffer (20 mM Tris-HCl, pH 7.4, 137 mM NaCl, 0.1% Tween 20) for 1 h. The membrane was then incubated with different antibodies diluted in TBS-T buffer or in TBS-T buffer containing 5% nonfat dry milk for 4 h at RT or overnight at 4ºC followed by three washes with TBS-T buffer for 10 min each. The membrane then was incubated with horseradish peroxidase-tagged secondary antibodies diluted in TBS-T buffer containing 5% nonfat dry milk for 1 to 2 h and washed three times with TBS-T buffer for 10 min each. Binding of the secondary antibody was detected using the Western Lightning Enhanced Chemiluminescent Reagent (Perkin Elmer, Wellesley, MA) as recommended by the manufacturer.

2.13 ANTIBODIES

GFP, Actin, AID and HA-specific antibodies were obtained from Clontech, Calbiochem, Cell Signaling and BioLegend respectively. SP-specific antibody was prepared by immunization of rabbits following standard protocols of Cocalico Biologicals (Reamstown, PA) [24]. MMTV Gag-specific monoclonal antibodies were kindly provided by Dr. Tanya Golovkina (University of Chicago)[192]. Incubation with all primary

antibodies was performed at 4^{o}C overnight, whereas secondary antibody incubations were performed for 1 h at RT.

2.14 STATISTICAL ANALYSIS

Statistical significance for luciferase assays was determined using the two-tailed Student's *t* test. Statistical differences between tumor development induced by the wild-type and mutant viruses and Kaplan-Meier survival curves were calculated using SPSS software. Differences in distribution of mutations/clone were assessed by non-parametric Mann-Whitney tests. Correlation analysis was performed by calculating Spearman's correlation coefficient. Mann-Whitney tests and Correlation analysis was performed using Graph Pad Prism5. A p - value of less than 0.05 was considered to be significant in all *t*-tests.

Chapter 3: MMTV Rem Antagonizes Activation-Induced Cytidine Deaminase (AID)

3.1 RATIONALE

Retroviruses such as HIV-1 and MuLV have been shown to encode for regulatory proteins that antagonize the APOBEC/Apobec family of restriction factors during their replication within T lymphocytes [151, 170, 194-198]. Previous studies have shown that MMTV requires B and T cells for efficient transmission to the mammary gland [25, 193]. MMTV elicits a Sag-mediated T-cell response, but manages to evade the immune system prior to inducing mammary carcinomas. MMTV infection of mice deficient for murine Apobec3 (mA3) revealed accelerated mammary tumorigenesis and increased viral loads relative to wild-type C57BL/6 mice [179]. Currently, no MMTV-encoded protein has been shown to antagonize mA3.

MMTV Rem protein is synthesized as a precursor at the endoplasmic reticulum (ER) membrane, and gets cleaved into an N-terminal signal peptide (SP) and a C-terminal protein (Rem-CT) [24]. Rem is a doubly spliced version of the singly spliced MMTV *env* mRNA, which also generates SP in the same open reading frame. SP has functional similarities to HIV-Rev [22, 23]. However, function of the full-length Rem protein is still unknown. In this study, I assessed the potential role of full-length Rem (or the cleavage product Rem-CT) as an accessory factor *in vivo*, to antagonize the murine Apobec family of cytidine deaminases.

3.2 RESULTS

3.2.1 MMTVs lacking Rem expression have reduced incidence and increased latency of mammary tumors.

MMTV-encoded Rem is synthesized as a precursor protein at the ER membrane [22]. The precursor Rem gets cleaved by the signal peptidase into an N-terminal SP, which is a Rev-like protein. SP is also generated from the Env protein [23, 24]. Another Rem cleavage product, Rem-CT, is a truncated version of the Env surface (SU) and transmembrane (TM) domains (Fig. 1.5). To confirm expression of Rem-CT, 293 cells were transfected with either untagged or tagged Rem-expression plasmids. Both plasmids yielded a cleaved C-terminal product that could be detected with Rem-CT antibody, although the tagged version had higher steady-state levels (Fig. 3.1).

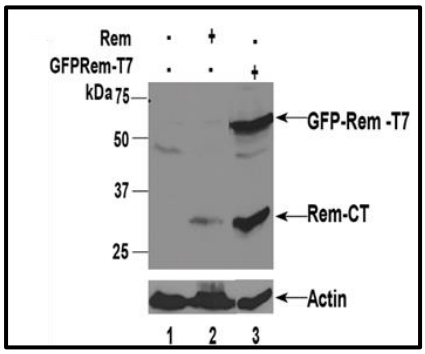

Figure 3.1: Western blotting of extracts from 293 cells transfected with untagged or tagged Rem expression constructs.

Positions of the tagged Rem precursor and cleaved Rem-CT (tagged and untagged) are shown in the upper panel after incubation of Western blots with Rem-CT-specific antibody. The bottom panel shows the same extracts with actin-specific antibody. This blot was performed by Nimita Halani.

To determine the function of Rem C-terminal sequences *in vivo*, an infectious MMTV provirus harboring a mutation within the *env* splice donor site (MMTV-SD) was used (Fig. 3.2). The SD mutation consists of 7 non-contiguous bases (GGGGTGAGT to GGACTCTCA) resulting in a single valine to leucine change in the Env protein. Previous studies have shown that this mutant produces normal viral RNA levels has and no replication defects in tissue culture [185]. Ablation of the SD2 site prevents expression of Rem, Rem-CT, as well as Sag from the intragenic *env* promoter. However, Sag expression from the upstream LTR promoter remains unaffected.

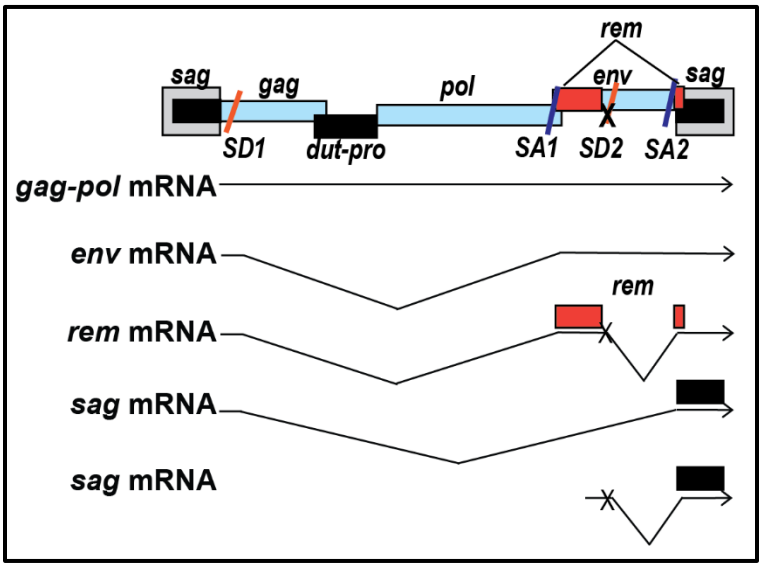

Figure 3.2: Strategy for generation of Rem-null MMTV provirus.

The thick gray boxes are long terminal repeats (LTRs); the thinner boxes show open reading frames. The *rem* coding region is shown in red. The SD2 mutation (designated by an X) eliminates the downstream SD site needed to generate the doubly spliced *rem* mRNA and the singly spliced *sag* mRNA from the internal *env* promoter [185].

Western blot analysis of XC (rat fibroblast) cells stably expressing MMTV-WT and MMTV-SD showed no difference in the levels of intracellular Gag protein expression (Fig. 3.3A). As expected RT-PCR for viral mRNAs indicated loss of *rem* upon introduction of the SD mutation (Fig. 3.3B). Additionally, to determine whether introduction of the SD2 mutation resulted in decreased SP activity, we co-transfected either MMTV-WT or MMTV-SD with pHM*Rluc* vector in XC cells. This vector has been shown to quantitatively measure the post-export activities of MMTV SP and to be dependent on the presence of a functional Rem-responsive element [23]. Similar levels of SP activity were observed for both wild-type and mutant proviruses (Fig. 3.4). Therefore, the loss of Rem expression has no effect on the level and processing of Env in tissue culture.

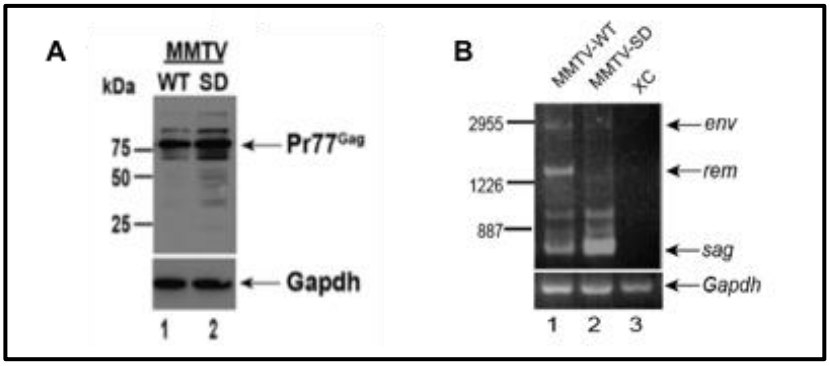

Figure 3.3: MMTV-WT and MMTV-SD produce equivalent amounts of Gag precursor in tissue culture cells & MMTV-SD proviruses lack *rem* mRNA synthesis.

A) Stably transfected XC rat cells expressing MMTV-WT or MMTV-SD were used to harvest cell extracts. Westerns were incubated with CA-specific (upper panel) or Gapdh-specific (lower panel) antibody. Western blotting with CA-specific antibody shows similar amounts of Gag precursor (Pr77) expression in cell extracts. B) RNA from XC fibroblast cells stably expressing MMTV-WT or MMTV-SD and from untransfected cells was used for RT-PCRs. Primers were designed for the MMTV LTR (upper panel) or *Gapdh* (lower panel).

To determine whether Rem C-terminal sequences serve an accessory function *in vivo*, BALB/cJ mice were injected intraperitoneally with $2x10^7$ XC cells expressing the MMTV-WT or MMTV-SD provirus. As expected from previously published experiments [185], MMTV-SD infection resulted in fewer mammary tumors with an increased latency compared to MMTV-WT (Fig 3.5).

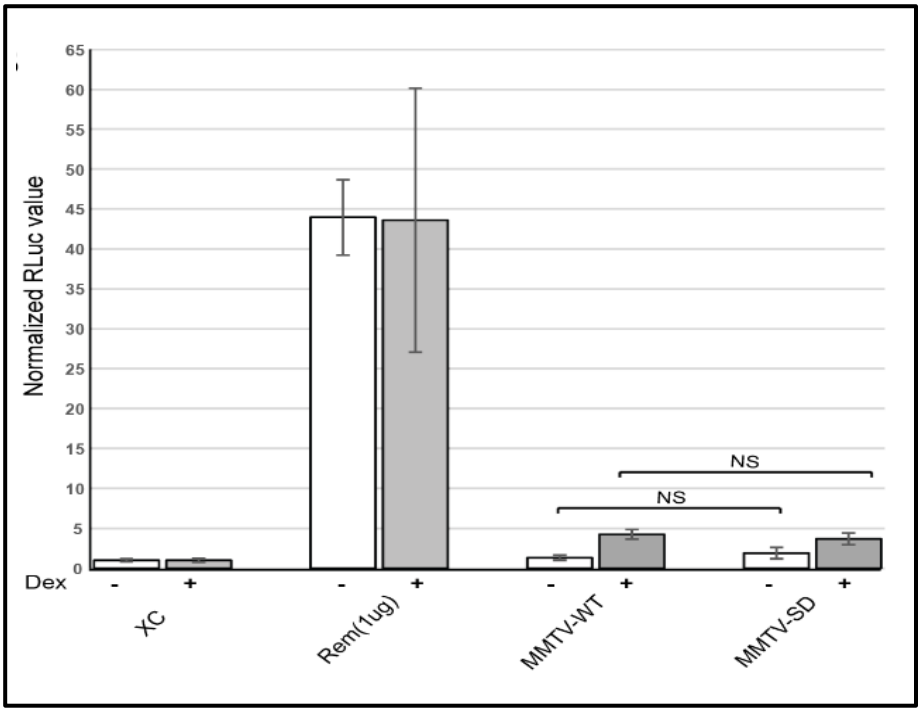

Figure 3.4: MMTV-WT and MMTV-SD have similar levels of SP activity.

Transient transfections of MMTV-WT and MMTV-SD were performed in XC rat fibroblasts. Cells were transfected and incubated in the presence or absence of dexamethasone (Dex) as indicated to stimulate the hormone-inducible MMTV LTR. SP activity as measured by a *Renilla* luciferase reporter assay is unaffected by the Rem-null MMTV provirus. NS = not significant. Transfection of the CMV-promoter-driven Rem expression plasmid was used as a positive control. The majority of SP activity is derived from Env cleavage, indicating no difference in Env production between MMTV-WT and MMTV-SD.

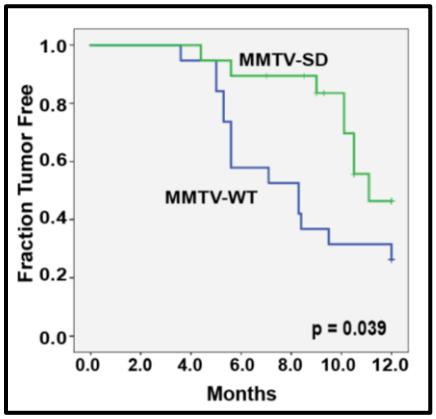

Figure 3.5: BALB/cJ mice infected with the Rem-null (MMTV-SD) virus develop mammary tumors with lower incidence and increased latency compared to tumors induced by MMTV-WT.

Kaplan-Meier plots reveal a difference in mammary tumor development between mice infected with MMTV-WT (blue line) or MMTV-SD (green line) (see p-value).

Next, to assess the role of Rem in restricting levels of MMTV infecting the mammary gland, proviral loads were determined in tumors obtained from wild-type and Rem-null virus-induced mammary tumors. Despite many attempts, none of the primer pairs tested could distinguish the infecting exogenous MMTV from the endogenous *Mtv*s from BALB/cJ mice in quantitative PCRs (data not shown). Therefore, semi-quantitative PCRs were performed on three independent tumors from different mice with primers for the MMTV polymerase region, with pol(4235+) and pol(5835-) and compared to DNA from uninfected BALB/cJ mice. The *Mtvr2* (Mouse mammary tumor virus receptor homolog 2) gene was used as a normalization control. Wild-type and Rem-null MMTV-induced tumors had increased proviral loads compared to DNA from uninfected mice (Fig. 3.6). Based on the ability of these primers to detect one diploid copy each of endogenous *Mtv*s *8* and *9* in

BALB/cJ mice [32], these results suggest that MMTV-induced tumors had acquired 4 to 7 haploid proviruses. Viral loads were higher in MMTV-WT induced tumors compared to those induced by MMTV-SD ($p<0.05$), consistent with restriction of viruses infecting the mammary gland as a consequence of the Rem-null mutation.

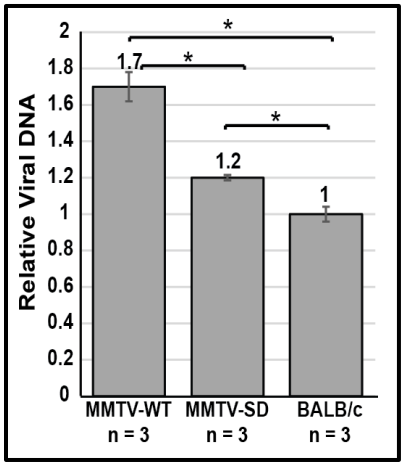

Figure 3.6: MMTV-SD-induced mammary tumors have reduced viral DNA levels relative to those in MMTV-WT-induced tumors.

PCRs were performed using DNA from three mammary tumors derived by inoculation of three independent BALB/cJ mice with either MMTV-WT or MMTV-SD and compared with uninfected BALB/cJ DNA containing endogenous *Mtv* proviruses. Statistical significance between columns is indicated by an asterisk ($p<0.05$).

3.2.2 Proviral clones from mammary tumors induced by MMTVs lacking Rem expression show increased viral genome mutation.

The cytidine deaminase mA3 has been shown to restrict MMTV infection *in vivo* [179], but none of the MMTV-encoded proteins are known to antagonize this enzyme. Since fewer tumors were induced by MMTV-SD and these tumors had a lower viral load,

I tested whether the presence of Rem facilitates MMTV pathogenesis by inhibiting the activity of mA3. If Rem is an mA3 antagonist, then the absence of Rem would be expected to increase G to A mutations on the viral genome plus strand.

DNA was extracted from five independent mammary tumors induced by MMTV-WT as well as five MMTV-SD-induced tumors. The MMTV envelope gene was amplified using virus specific primers (env 7254+ and env 420-). Individual clones were obtained for Sanger sequencing, which allows analysis of longer regions of the viral genome that will distinguish between endogenous *Mtv*s and exogenous virus integrations. Longer reads were also suitable for differentiating between proviral clones that had retained the original mutation in the SD site relative to proviruses that had lost the mutation by recombination with endogenous *Mtv*s as observed previously [185]. Since analysis of the proviral loads indicated that ~ 4 to 7 copies of the virus were acquired, only five clones from each mammary tumor were sequenced to avoid duplicate sampling. Sequencing results showed that the plus strand of the proviral clones obtained from MMTV-SD induced tumors had increased transition as well transversion mutations, compared to clones from tumors induced by MMTV-WT. The largest increase consisted of C to T transitions on the plus strand of the proviruses. Previous reports have shown that mA3-induced cytidine deamination of the proviral minus strand during reverse transcription results in G to A mutation of the viral plus strand [97, 151] (Table 3.1).

Since non-G to A mutations on the plus strand are not typically observed as a result of mA3 activity, the involvement of additional cytidine deaminases was analyzed. Replication in B cells is an absolute requirement for efficient transmission of MMTV to

the mammary gland. The cytidine deaminase AID is preferentially expressed in germinal center B cells [199] and has been shown to antagonize retrotransposon insertions [132]. AID-induced cytidine deamination is often followed by error-prone DNA repair activity resulting in additional transition and transversion mutations [200,201]. Thus, mutations observed in MMTV proviruses were analyzed for the WRC motif, reported to be a "hot-spot" for AID-induced base changes within immunoglobulin gene variable regions [202].

The frequency of cytidine mutations in the WRC context was almost five-fold higher in the Rem-null MMTV-SD-induced tumors (Table 3.1). The frequency of cytidine mutations in the SYC context, which has been described as an AID "cold spot" and less frequently mutated [203,204], were also higher in the proviruses obtained from MMTV-SD tumors. However, the fold-increase was lower than that observed for the WRC motif. Cytidine changes in two other sequence contexts (TYC and ATC) also were analyzed. Analysis of the frequency of changes in the TYC motif, typical of those induced by mA3 [110, 171], showed mutations only in the proviruses obtained from the MMTV-SD-induced tumors. TYC mutations were not detected in MMTV-WT proviruses. Mutations in the ATC context also have been associated with mA3 *in vitro* [180], and analysis of these mutations indicated a slight increase in MMTV-SD proviruses compared to MMTV-WT proviruses. Because of the wide variability in number of mutations/clone, I could not analyze this tabular data statistically (see below). Nevertheless, these data are consistent with increased mutations by mAID and/or mA3 after loss of the SD site.

Table 3.1: Mutation frequency in MMTV-WT and MMTV-SD proviruses from BALB/cJ mammary tumors by Sanger sequencing.

Mutation	MMTV-WT Mutation Frequency[1]	MMTV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	0.36	1.28	3.6
A to G	0.56	1.22	2.2
C to T	**0.08**	**0.66**	**8.3**
T to C	1.08	1.38	1.3
Transition	2.08	4.54	2.2
Transversion	0.56	1.00	1.8
WRC	**0.12**	**0.59**	**4.9**
SYC	0.08	0.28	3.5
TYC	0.00	0.53	-
ATC	0.12	0.19	1.6

[1] Mutations/number of clones. Based on Sanger sequencing of 1,100 bp of the plus strand of the proviral *env* gene clones obtained from five independent BALB/cJ tumors induced by MMTV-WT (n=25) or MMTV-SD (n=32).

Interestingly, about 22% of the clones obtained from MMTV-SD-induced tumors showed loss of the introduced SD mutation as noted in a previous study [185]. All of these clones had acquired the SD site sequence exactly matching that of the wild-type virus (GGACTCTCA to GGGGTGAGT), suggesting selection pressure for splice site restoration *in vivo*. Because of the number of base mutations, reversion of the SD site mutation is most likely a result of a recombination event between the endogenous *Mtv*s and exogenous virus, as reported previously in MMTV-induced mammary tumors [28, 185, 205]. MMTV-Rem protein had not been discovered when such recombination events were first reported. Hence, occurrence of this event had been attributed solely to the requirement of Sag expression from the intragenic envelope promoter (see Fig. 3.2).

Since a large variability in the number of mutations per clone was observed, a non-parametric *t*-test (Mann-Whitney) was employed to determine statistically significant differences in the distribution of mutations/clone across different groups. Proviruses with reverted SD-sites had higher numbers of WRC and ATC motif mutations compared to proviruses with SD-site mutations that had not undergone recombination with endogenous *Mtv*s (Fig. 3.7A and D). In contrast, mutations within SYC and TYC motifs were not significantly different regardless of the SD-site reversion (Fig. 3.7B and C). Despite having higher mutations, MMTV-SD recombinant clones lacked stop codons within the *env* open reading frame. This result further suggests that selection of replication-competent virus during MMTV pathogenesis *in vivo* (Fig. 3.8) results in altered *sag* and/or *rem* expression. Therefore, these data indicate that reduction of one or both of these genes through the SD mutation reduces the ability of MMTV to induce mammary tumors as well as the proviral load and genomic integrity of proviruses in these tumors.

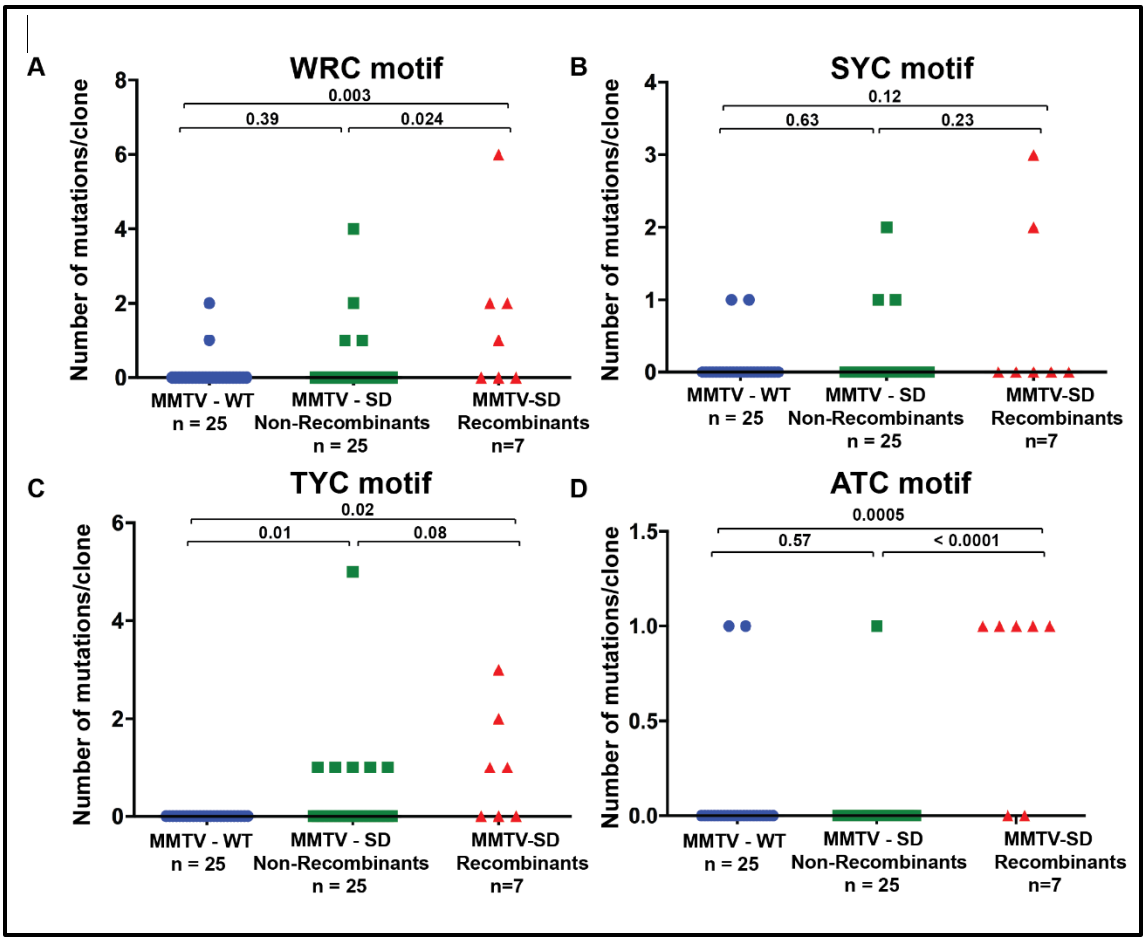

Figure 3.7: Mutational analysis of the proviral envelope gene from BALB/cJ mammary tumors induced by MMTV-WT or MMTV-SD.

Number of C mutations within different motifs on either proviral strand is given for each clone. Number of clones (n) is indicated. Sequences were obtained from independent clones from five tumors in different animals. (**A**) Number of mutations/clone in the WRC motif typical of AID mutation hotspots. (**B**) Number of mutations/clone in the SYC motif. (**C**) Number of mutations/clone in the TYC motif typical of mA3 mutation hotspots. (**D**) Number of mutations/clone in the ATC motif. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

```
MMTV-ENV        AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYTPPVNKEVHRWYEAGW   314
Revertant1      AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYTPPVNKEVHRWYEAGW   314
Revertant2      AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYTPPVNKEVHQWYEAGW   314
Revertant3      AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYTPPVNKEVHQWYEAGW   314
Revertant4      AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYTPPVNKEVHRWYEAGW   314
Revertant5      AFPDQGVSFSPKGALGLLWDFSLPSPNVDQSEQIKNKKDLFGNYTPPVNKEVHRWYEAGW   314
Revertant6      AFPDQGVSFSPKGALGLLWDFSLPSPNIDQSEQIKNKKDLFGNYTPPVNKEVHRWYEAGW   314
Revertant7      AFPDQGVSFSPKGALGLLWDFSLPSPNIDQSEQIKNKKDLFGNYTPPVNKEVHRWYEAGW   314

MMTV-ENV        VEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMIPTSACVTYP   374
Revertant1      VEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMIPTSACVTYP   374
Revertant2      VEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMIPTSACVTYP   374
Revertant3      VEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMIPTSACVTYP   374
Revertant4      VEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMIPTSACVTYP   374
Revertant5      VEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRHLILKRPGFQEHEMISTSACVTYP   374
Revertant6      VEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRHLILKRPGFQEHEMISTSACVTYP   374
Revertant7      VEPTWFWENSPKDPDDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMIPTSACVTYP   374

MMTV-ENV        HAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYVLLPVDIGDE   434
Revertant1      YAILLGLPQLIDIEKRGSTFHTSCSSCRLTNCLDSSAYDYAAIIVKRPPYVLLPVDIGDE   434
Revertant2      YAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYVLLPVDIGDE   434
Revertant3      YAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYVLLPVDIGDE   434
Revertant4      YAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYVLLPVDIGDE   434
Revertant5      YAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYVLLPVDIGDE   434
Revertant6      YAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSFAYDYAAIIVKRPPYVLLPVDIGDE   434
Revertant7      YAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYVLLPVDIGDE   434

MMTV-ENV        PWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494
Revertant1      PWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494
Revertant2      PWFDDSAIQTFKYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494
Revertant3      PWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494
Revertant4      PWFDDSAIQTFRYATDLIRAKRFVTAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494
Revertant5      PWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494
Revertant6      PWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494
Revertant7      PWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKEMQTATFVNN   494

MMTV-ENV        LHRNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHANYDFICVTPL   554
Revertant1      LHRNVTLALSEQRIIDLKLEARLNALEEVVLESGQDVANLKTRMSTRCHANYDFICVTPL   554
Revertant2      LHKNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHANYDFICVTPL   554
Revertant3      LHKNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHANYDFICVTPL   554
Revertant4      LHRNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHANYDFICVTPL   554
Revertant5      LHRNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHANYDFICVTPL   554
Revertant6      LHRNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHANYDFICVTPL   554
Revertant7      LHRNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHANYDFICVTPL   554

MMTV-ENV        PYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDTVDLSGLAQS         608
Revertant1      PYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDTVDLSGLAQS         608
Revertant2      PYNASESWERTKAHLLGIWNDNEISYNIQELTNPISDMSKQHIDTVDLSGLAQS         608
Revertant3      PYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDAVDLSGLAQS         608
Revertant4      PYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDAVDLSGLAQS         608
Revertant5      PYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDTVDLSGLAQS         608
Revertant6      PYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKRHIDTVDLSGLAQS         608
Revertant7      PYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDTVDLSGLAQS         608
```

Figure 3.8: MMTV-SD recombinant proviruses lack stop codons.

Amino acid differences between the MMTV-WT and MMTV-SD recombinant proviruses in the *env* gene as determined for individual clones by Sanger sequencing. Increased viral gene mutations in the absence of Rem yield recombinants that lack stop codons within the *env* open reading frame of the MMTV-SD-recombinant proviruses.

3.2.3 Viral genome hypermutation of Rem-null virus is also observed in tumors induced by the Sag-independent variant TBLV.

Although the *sag* transcript may be generated from the upstream LTR promoter in the SD site mutant MMTV [185], increased mutation of the viral genome could be attributed to loss of Sag expression from the *env* promoter or from loss of Rem expression. To distinguish between the roles of these two accessory proteins, the thymotropic variant TBLV was used for *in vivo* experiments. TBLV is a naturally occurring MMTV strain that encodes a truncated non-functional *sag* gene [64, 65, 70]. Despite this truncation TBLV can cause lymphomas in mice. Similar to results with MMTV-WT and MMTV-SD, SP activity was not reduced in Rem-null TBLV proviruses (TBLV-SD) whereas expression of *rem* mRNA was blocked (Fig. 3.9A and 3.9B). Also, no difference in virus production was observed between cells expressing TBLV-WT or TBLV-SD (Fig. 3.10).

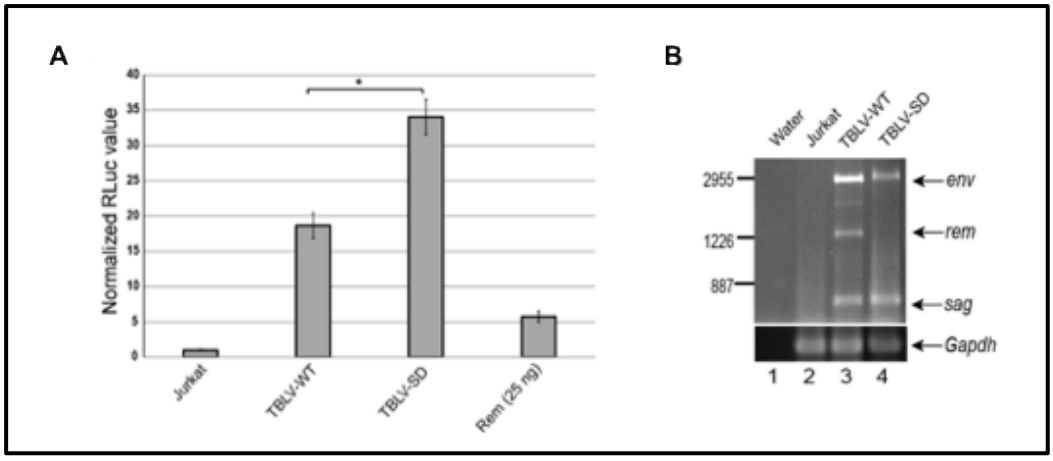

Figure 3.9: TBLV-WT and TBLV-SD have similar levels of SP activity & TBLV-SD proviruses lack *rem* mRNA synthesis.

A)Transient transfections of TBLV-WT and TBLV-SD in human Jurkat cells. SP activity as measured by a *Renilla* luciferase reporter assay is preserved by the Rem-null TBLV provirus. Transfection of CMV-promoter-driven Rem expression plasmid was used as a positive control. The asterisk indicates significance ($p<0.05$). **B)** RNA from human Jurkat cells stably expressing either TBLV-WT or TBLV-SD and from untransfected cells was used for RT-PCRs. Primers were designed for the TBLV LTR (upper panel) or *Gapdh* (lower panel). These experiments were performed by Almas Ali.

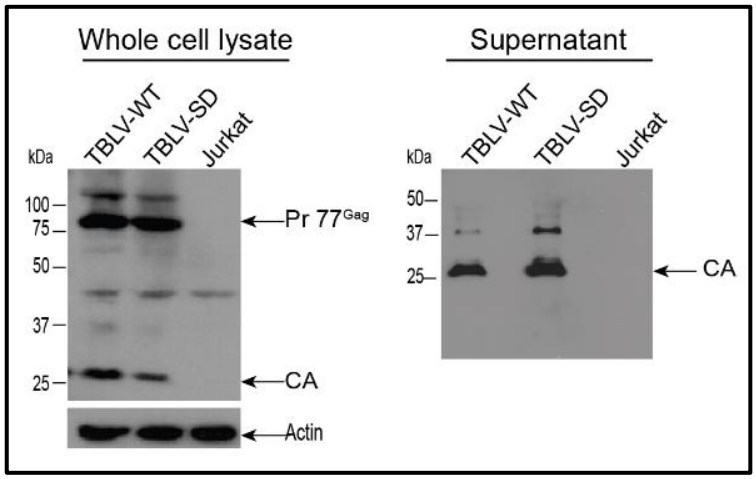

Figure 3.10: TBLV-WT and TBLV-SD produce equivalent amounts of virus in tissue culture.

Stably transfected Jurkat cells expressing TBLV-WT or TBLV-SD were used to harvest supernatants or for preparation of cytosolic extracts. Western blotting shows similar amounts of Gag expression in whole-cell lysates (left panel) and released virus in supernatants **(**right panel**)** from cells transfected with expression plasmids for TBLV-WT or TBLV-SD. Western blots were incubated with CA-specific or actin-specific antibody. Arrows indicate the positions of cleaved capsid (CA) or capsid precursor ($Pr77^{Gag}$).

Thus, to confirm the role of Rem in preventing increased mutation of the viral genome, BALB/cJ mice were injected intraperitoneally with $2x10^7$ Jurkat cells stably expressing TBLV-WT or TBLV-SD. Contrary to MMTV-SD, Rem-null TBLV did not result in reduced or delayed tumors (Fig. 3.11A). However higher proviral loads were maintained in TBLV-WT induced tumors compared to TBLV-SD tumors (Fig. 3.11B). This result suggests that differences in tumor incidence and latency are a Sag-dependent phenomenon as reported previously [185]. Rem C-terminal sequences on the other hand may play a role at an earlier step during the virus life cycle, by antagonizing Apobec

enzymes in the lymphoid population prior to transmission of the virus to the mammary gland [206].

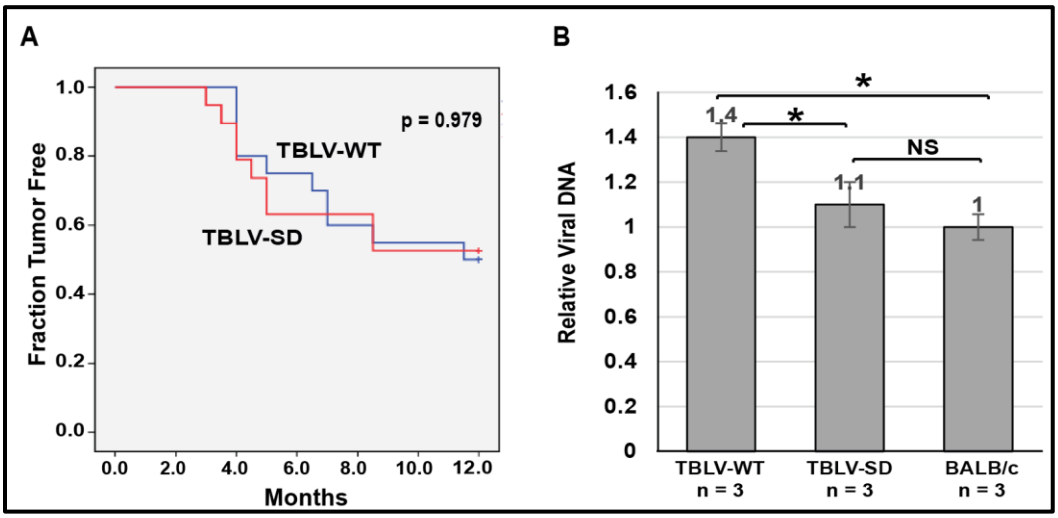

Figure 3.11: TBLV-SD-induced thymic tumors have reduced viral DNA levels relative to TBLV-WT-induced tumors.

(**A**) Kaplan-Meier plots for T-cell lymphoma development in BALB/cJ mice injected with TBLV-WT (blue line) or TBLV-SD (red line) are shown with p-value. (**B**) TBLV-SD-induced thymic tumors have reduced viral DNA levels relative to TBLV-WT-induced tumors. PCRs were performed using DNA from three thymic tumors derived by inoculation of three independent BALB/cJ mice with either TBLV-WT or TBLV-SD. PCR product levels were compared with uninfected BALB/cJ DNA containing endogenous *Mtv* proviruses. Statistical significance between columns is indicated by an asterisk ($p<0.05$). NS = not significant.

To determine whether the absence of Rem results in increased proviral mutations, DNA was extracted from thymic tumors induced by wild-type or mutant TBLV. The *env* gene was amplified by PCR using virus-specific primers. Three independent TBLV-WT tumors as well as TBLV-SD-induced tumors were analyzed. After cloning of PCR

products, sequencing analysis revealed that the TBLV genome was also subjected to higher mutations in the absence of Rem. The C to T transitions showed the highest increase between the wild-type versus SD mutant proviruses, although the G to A transitions had the highest frequency (Table 3.2) as observed for MMTV. The increased frequency of mutations was not observed for the cellular gene *c-Myc*, which also was amplified from TBLV-WT and TBLV-SD-induced tumors (Table 3.3).

Table 3.2: Mutation frequency in TBLV-WT and TBLV-SD proviruses from BALB/cJ thymic tumors by Sanger sequencing.

Mutation	TBLV-WT Mutation Frequency[1]	TBLV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	2.02	2.98	1.5
A to G	0.60	1.27	2.1
C to T	**0.08**	**0.73**	**9.1**
T to C	0.86	1.53	1.8
Transition	3.56	6.51	1.8
Transversion	0.10	0.71	7.1
WRC	**0.26**	**0.67**	**2.6**
SYC	0.14	0.69	4.9
TYC	0.84	1.65	2.0
ATC	0.18	0.47	2.6

[1] Mutations/number of clones. Based on Sanger sequencing of 1,100 bp of the plus strand of the proviral *env* gene clones obtained from three independent BALB/cJ tumors induced by TBLV-WT (n=50) or TBLV-SD (n=55).

Table 3.3: Mutation frequency in the *c-Myc* gene from TBLV-induced BALB/cJ thymic tumors by Sanger sequencing.

Mutation	TBLV-WT Mutation Frequency[1]	TBLV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	0.10	0.10	1.0
A to G	0.30	0.30	1.0
C to T	0.10	0.10	1.0
T to C	0.20	0.20	1.0
Transversions	0.10	0.02	0.2

[1] Mutations/number of clones. Based on Sanger sequencing of 720 bp of the plus-strand of the *c-Myc* gene clones (~40) obtained from three independent BALB/cJ T-cell tumors induced by TBLV-WT or TBLV-SD.

Since unlike MMTV-induced tumors, TBLV induces polyclonal tumors, high-throughput sequencing was used also for the analysis of proviruses. Averaged reads also revealed that TBLV-SD proviruses had numerous G to A and C to T as well as other transition mutations relative to TBLV-WT proviruses (Table 3.4A). A similar result was observed for a portion of the TBLV polymerase gene (Table 3.4B), but not cellular *Gapdh* gene (Tables 3.4C).

Table 3.4: Averaged mutations from proviral and cellular genes from TBLV-induced BALB/cJ thymic tumors by Illumina sequencing.

Table 3.4A. Averaged mutations in the TBLV-WT versus TBLV-SD proviral envelope genes from T-cell tumors by Illumina sequencing			
Mutation	**TBLV-WT Tumors[1]**	**TBLV-SD Tumors[1]**	**Fold Increase**
G to A	12	52	4.3
A to G	2	17	8.5
C to T	0	14	>14.0
T to C	4	12	3.0
Transversions	2	4	2.0
Table 3.4B. Averaged mutations in the TBLV-WT versus TBLV-SD proviral polymerase genes from T-cell tumors by Illumina sequencing			
Mutation	**TBLV-WT Tumors[2]**	**TBLV-SD Tumors[2]**	**Fold Increase**
G to A	13	35	2.7
A to G	3	15	5.0
C to T	1	15	15.0
T to C	1	16	16.0
Table 3.4C. Averaged mutations in the *Gapdh* gene from TBLV-WT and TBLV-SD-induced T-cell tumors by Illumina sequencing			
Mutation	**TBLV-WT Tumors[3]**	**TBLV-SD Tumors[3]**	**Fold Increase**
G to A	15	15	1.0
A to G	1	1	1.0
C to T	17	17	1.0
T to C	4	4	1.0
Transversions	3	3	1.0

[1] Number of mutations in ~60,000 reads averaged at each base within 1,371 bp of the plus-strand of proviral envelope genes in three independent BALB/cJ tumors induced by TBLV-WT or TBLV-SD using Illumina sequencing of PCR products. Mutations were identified as ≥ 2% different from the reference sequence.

[2] Number of mutations in ~60,000 reads averaged at each base within 1,619 bp of the plus-strand of the proviral polymerase genes in three independent BALB/cJ tumors induced by TBLV-WT or TBLV-SD using Illumina sequencing of PCR products. Mutations were identified as those that were ≥ 2% different from the reference sequence. No transversions were observed.

[3] Number of mutations in ~60,000 reads averaged at each base within 581 bp of the plus-strand of the *Gapdh* genes in three independent BALB/cJ tumors induced by TBLV-WT or TBLV-SD using Illumina sequencing of PCR products. Mutations were identified as those that were ≥ 2% different from the reference sequence.

To determine whether SD site reversion occurred in T-cell tumors, I also analyzed the frequency of this event in TBLV-SD proviruses. Although the frequency of SD reversion was lower (9%) than in MMTV-SD proviruses (22%), both types of reversion were consistent with recombination with endogenous *Mtv*s [185], i.e., all 6 of the introduced nucleotide mutations were corrected to the sequence of the endogenous *Mtv*s *Mtv8* or *Mtv9* proviruses found in BALB/cJ mice [32]. In contrast to MMTV-SD proviruses from mammary tumors, 80% of the TBLV-SD recombinants had stop codons in the *env* gene (Fig. 3.12). Since TBLV is an end-point virus that is not transmitted to the mammary gland, this lymphomagenic virus may be subjected to a lower selective pressure during replication in lymphoid cells. Nevertheless, selection for reversion of the TBLV-SD site occurs within lymphoid cells. More importantly, since TBLV is a Sag-independent virus, SD reversion points towards a requirement of Rem protein to allow efficient virus replication in lymphocytes *in vivo*.

```
TBLV-ENV                AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKSKKDLFGNYTPPVNKE        305
Revertant1              AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKNKKDLFGNYTPPVNKE        305
Revertant2              AFPDQGVSFSPKGALGLLWDFSLPSPSVDQSDQIKNKKDLFGNYTPPVNKE        305
Revertant3              AFPDQGVSFSPKGALGLLWDFSLPSPNIDQSEQIKNKKDLFGNYTPPVNKE        305
Revertant4              AFPDQGVSFSPKGALGLLWDFSLPSPNIDQSEQIKNKKDLFGNYTPPVNKE        305
Revertant5              AFPDQGVSFSPKGALGLLWDFSLPSPNIDQSEQIKNKKDLFGNYTPPVNKE        305

TBLV-ENV        VHRWYEAGWVEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMI        365
Revertant1      VHRWYEAGWVEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRHLILKKPGFQEHEMI        365
Revertant2      VHRWYEAGWVEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRHLILKKPGFQEHEMI        365
Revertant3      VHRWHEAGWVEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRYLILKRPGFQEHDMI        365
Revertant4      VHRWYEAGWVEPTWLWENSPKDPNDRDFTALVPHTELFRLVAAPRHLTLKRPGFQEHEMI        365
Revertant5      VHRWYEAGWVEPTWFWENSPKDPNDRDFTALVPHTELFRLVAASRHLILKRPGFQEHEMI        365

TBLV-ENV        PTSACVTYPHAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYV        426
Revertant1      PTSACVTYPYAILLGLPQLIDIEKRGSTFHISCSSRRLTNCLDSSAYDYAAIIVKRPPYV        426
Revertant2      PTSACVTYPYAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRPPYV        426
Revertant3      PTSACVTYPYAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSSAYDYAAIIVKRLPYV        426
Revertant4      STSACVTYPYAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSFAYDYAAIIVKRPPYV        426
Revertant5      STSACVTYPYAILLGLPQLIDIEKRGSTFHISCSSCRLTNCLDSFAYDYAAIIVKRPPYV        426

TBLV-ENV        LLPVDIGDEPWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKE        486
Revertant1      LLPVDIGDEPWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKE        486
Revertant2      LLPVDIGDEPWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIATITSFAVATTALVKE        486
Revertant3      LLPVDIGDEPWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALAKE        486
Revertant4      LLPVDIGDEPWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKE        486
Revertant5      LLPVDIGDEPWFDDSAIQTFRYATDLIRAKRFVAAIILGISALIAIITSFAVATTALVKE        486

TBLV-ENV        MQTATFVNNLHRNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHAN        546
Revertant1      MQTATFVNNLHRNVTLALSEQRIIDLKLKARLNALEEVVLELGQDVANLKTRMSTRCHAN        546
Revertant2      MQAATFVNNLHRNVTLALSEQRIIDLKLKARLNALEEVVLELGQDVANLKTRMSTRCHAN        546
Revertant3      MQTATFVNNLHRNVTLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHAN        546
Revertant4      MQTATFVNNLHRNITLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHAN        546
Revertant5      MQTATFVNNLHRNITLALSEQRIIDLKLEARLNALEEVVLELGQDVANLKTRMSTRCHAN        546

TBLV-ENV        YDFICVTPLPYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDTVDLSGL        606
Revertant1      YDFICVTPLPYNASESWERTRAHLLGIWNDNEISYNIRELTNLISDMSKQHIDAADLSGL        606
Revertant2      YDFICVTPLPYNASESWERTRAHLLGIWNDNEISYNIQELTNLISDMSKQHIDAVDLSGL        606
Revertant3      YDFICVTPLPYNASESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDTVDLSGL        606
Revertant4      YDFICVTPLPYNVSESWERTKAHLLGIWNDNEISYNIQELTNLISDMSKQHIDTVDLSGL        606
Revertant5      YDFICVTPLPYNASGSWERTKAHLLGI*NDNEISYNIQELTNLISDMSKQHIDTVDLSGL        606

TBLV-ENV        AQSFANGVKALNPLDWTQY                                                 625
Revertant1      AQSFANGVKALNPLD*TQY                                                 625
Revertant2      AQSFANGVKALNPLD*TQY                                                 625
Revertant3      AQSFANGVKALNPLD*TQY                                                 625
Revertant4      AQSFANGVKALNPLDWTQY                                                 625
Revertant5      AQSFANGVKALNPLNWTQY                                                 625
```

Figure 3.12: TBLV-SD recombinant proviruses have stop codons.

Amino acid differences between TBLV-WT and TBLV-SD- recombinant proviruses in the *env* gene as determined for individual clones by Sanger sequencing. Unlike MMTV-SD recombinants, increased viral gene mutations in the absence of Rem result in increased stop codons within the *env* open reading frame of the TBLV-SD-recombinant provirus. Stop codons are shown with an asterisk and are boxed in red.

To determine the specificity of other mutations within the TBLV *env* gene, I analyzed proviral mutations occurring in motifs recognized by various Apobec enzymes. The distribution of mutations/clones was significantly different between TBLV-SD mutants and recombinants in the WRC, SYC and TYC motifs, but not the ATC motif. (Fig. 3.13). Both MMTV and TBLV proviral clones showed increased TYC but not WRC motif mutations, regardless of the SD-site reversion. These results suggest that these mutations are a result of the activity of distinct enzymes, perhaps in different cell types. Correlation analysis was also conducted to examine the relationship between the SD-site reversion in MMTV and TBLV proviruses and the number of mutations in different motifs. Calculation of Spearman's coefficient revealed a significant correlation between the number of recombinants and mutations in each of the motifs (Fig. 3.14). In addition, *Mtv8* and *Mtv9* are known to be expressed in lymphocytes, suggesting that recombination and mutagenesis occurs in these cells [27]. Since TBLV is not transmitted to the mammary gland, these data are consistent with multiple Apobec-type mutations, including those induced by AID and mA3, occurring during viral replication in hematopoietic cells [70].

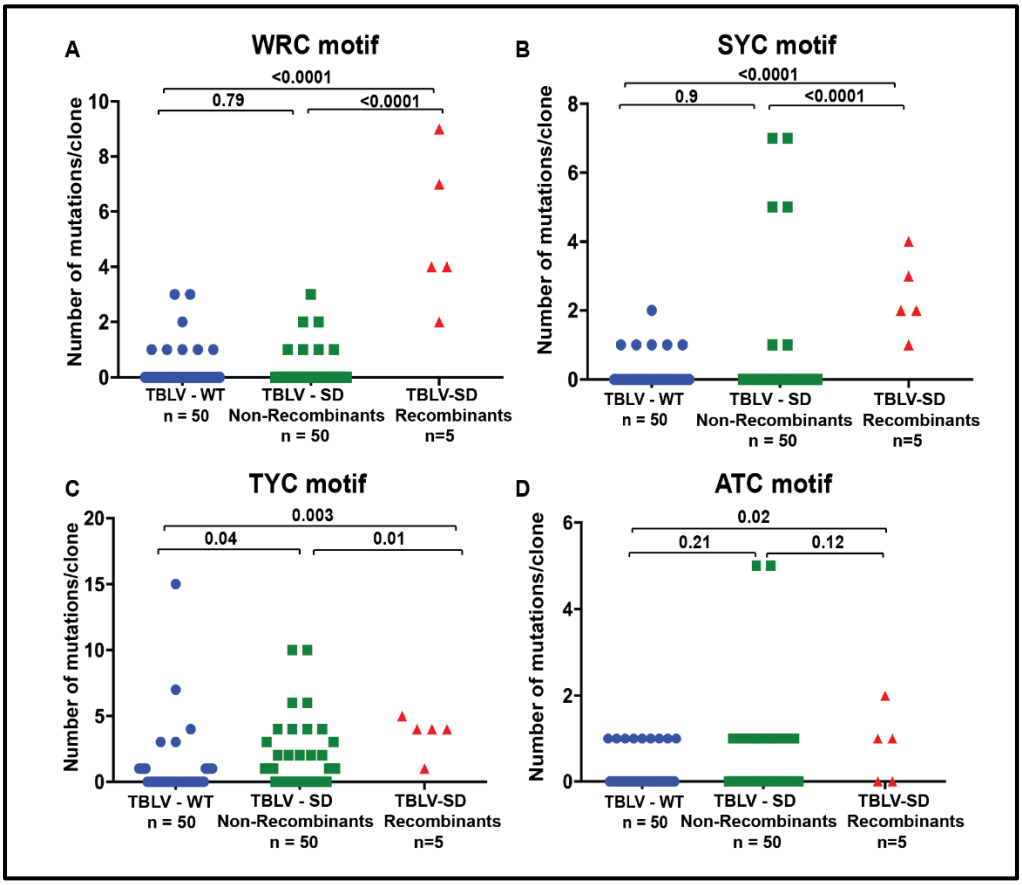

Figure 3.13: Mutational analysis of the proviral envelope gene from BALB/cJ thymic tumors induced by TBLV-WT or TBLV-SD.

Number of C mutations within different motifs on either proviral strand is given for each clone. Number of clones analyzed (n) is indicated. Sequences were obtained from independent clones from five tumors in different animals. (**A**) Number of mutations/clone in the WRC motif typical of AID mutation hotspots. (**B**) Number of mutations/clone in the SYC motif. (**C**) Number of mutations/clone in the TYC motif typical of mA3 mutation hotspots. (**D**) Number of mutations/clone in the ATC motif. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

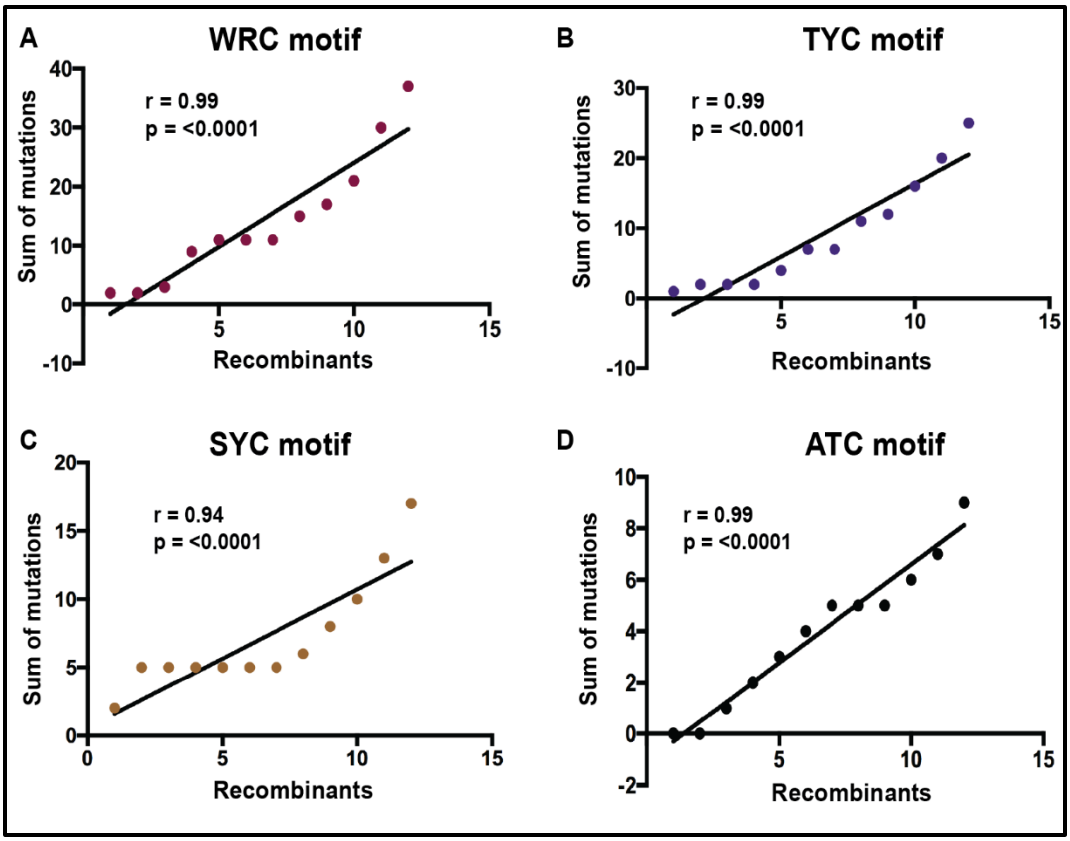

Figure 3.14: Correlation analysis of the number of recombinants and sum of mutations in different motifs from MMTV and TBLV-induced tumors.

(A-D) Spearman's correlation coefficient (r) indicating high correlation between number of recombinants and sum of mutations in WRC, TYC, SYC and ATC motifs, respectively. Correlation coefficient p-values are shown on the graph.

Several attempts also were made to study AID-induced mutations of the viral genome in tissue culture. Jurkat cells stably producing TBLV-WT or TBLV-SD were transduced with murine leukemia virus vectors expressing murine AID (pMX-AID). Transduced cells were grown in puromycin to select for AID-positive cells and DNA was extracted after ten days to analyze for mutations within the TBLV *env* gene. PCRs were performed with TBLV *env*-specific primers, and products were cloned. Sequencing analysis of individual clones revealed no difference in mutations in the absence and

presence of Rem expression (Table 3.5). Because Jurkat cells are of human origin and do not express the receptor for TBLV [64], these cells release virus after transfection, but cannot be re-infected. This result suggests that the increased mutations observed in the *in vivo* studies do not occur in cells releasing viral particles, but in the target cells following reverse transcription.

Table 3.5: Mutation frequency in TBLV-WT and TBLV-SD proviruses from AID-expressing Jurkat cells by Sanger sequencing.

	pMX-AID TBLV-WT Mutation Frequency[1]	pMX-AID TBLV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	1.20	0.90	0.8
A to G	0.70	0.70	1.0
C to T	0.150	0.13	0.9
T to C	1.10	1.10	1.0
Total	3.20	2.80	0.9

[1]Mutations/number of clones. Based on Sanger sequencing of 1,100 bp of the plus-strand of the proviral *env* gene clones obtained from pMX-AID and TBLV expressing Jurkat cells (TBLV-WT (n=34) or TBLV-SD (n=32).

3.2.4 Most viral genome hypermutation in tumors induced by Rem-null virus is abolished in AID-KO BALB/cJ mice.

To further define the involvement of AID in restricting MMTV infection, AID-KO BALB/cJ mice were injected with $2x10^7$ XC cells expressing MMTV-WT and MMTV-SD. These mice were generated by crossing AID-KO C57BL/6 mice from Dr. Honjo [93] to BALB/cJ mice, genotyping to select for AID-KO litter and followed by ten similar backcrosses with BALB/cJ. Although the appearance of mammary tumors was slightly delayed, no difference in tumor incidence or latency was observed upon infection of

BALB/cJ AID-KO mice with the Rem-null virus as opposed to results from MMTV infection of wild-type BALB/cJ mice (Figs. 3.15A, 3.16). Additionally, unlike infection of wild-type BALB/cJ mice with wild-type and Rem-null MMTV, analysis of proviral loads did not yield a significant difference between MMTV-WT and MMTV-SD-induced tumors in AID-KO BALB/cJ mice (Fig. 3.15B). This result strengthens the conclusion that differences in viral load are a Rem-dependent phenomenon whereas reduction in tumor incidence is a Sag-mediated effect.

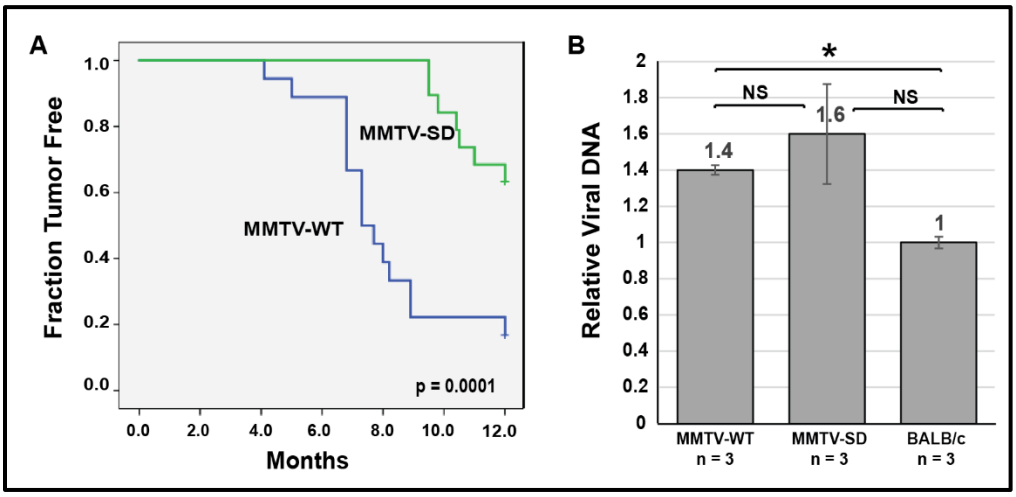

Figure 3.15: MMTV-WT and SD-induced mammary tumors in BALB/cJ AID-KO mice have different incidences and latencies, but similar viral loads.

(**A**) Kaplan-Meier plots for mammary tumor development in AID-KO BALB/cJ mice injected with MMTV-WT (blue line) or MMTV-SD (green line) are shown with p-value. (**B**) Proviral loads in MMTV-WT and MMTV-SD-induced mammary tumors relative to endogenous *Mtv* proviruses in uninfected AID-KO mice. PCRs were performed using DNA from three mammary tumors derived by inoculation of three independent BALB/cJ AID-KO mice with either MMTV-WT or MMTV-SD. Results were compared with uninfected BALB/cJ DNA containing endogenous *Mtv* proviruses. Statistical significance between columns is indicated by an asterisk ($p<0.05$). NS = not significant. Although the values were not significantly different (NS), the trend indicated that MMTV-SD proviral loads were higher than the endogenous *Mtv* levels.

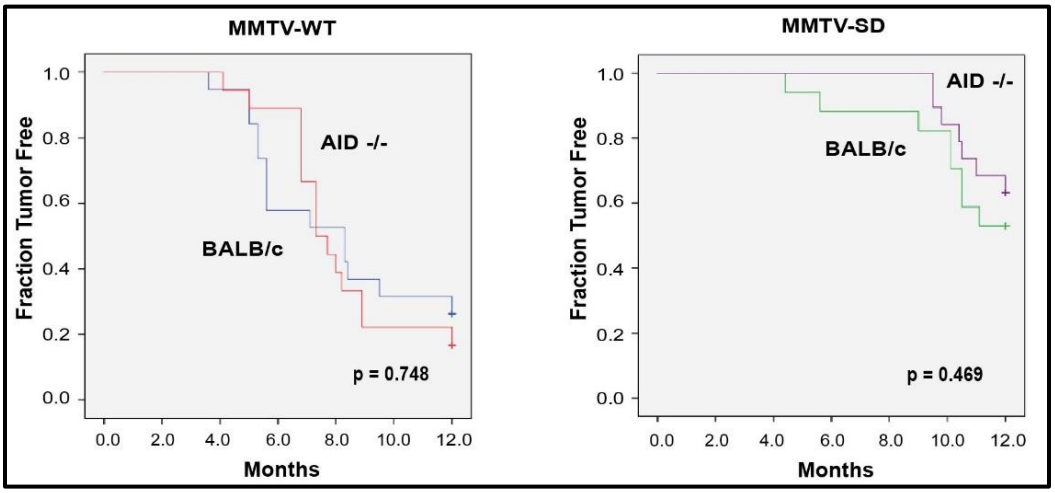

Figure 3.16: The incidence and latency of MMTV-induced mammary tumors does not differ between wild-type and AID-KO BALB/cJ mice.

Kaplan-Meier survival plots for MMTV-WT-infected wild-type BALB/cJ (blue) and AID-KO (red) mice (left panel) as well as MMTV-SD-infected wild-type BALB/cJ (green) and AID-KO (purple) mice (right panel) showed no significant difference (p-values are shown).

To determine the effect of AID ablation on proviral mutagenesis, proviruses were analyzed from five independent MMTV-WT and MMTV-SD tumors. Total cellular DNA was extracted, cloned with *env*-specific primers, and the resulting clones were subjected to Sanger sequencing. Although G to A mutations showed an increase in MMTV-SD compared to MMTV-WT proviruses, C to T mutations did not show the increase observed in wild-type BALB/cJ (Table 3.6 vs Table 3.1). As expected, the increase in AID-specific WRC motif mutations was abolished. Surprisingly, the fold-increase in TYC motifs typical of mA3-induced mutations was also abolished. Further, SYC motifs described as AID "cold-spots" [203, 204], showed the highest fold increase.

Table 3.6: Mutation frequency in MMTV-WT and MMTV-SD proviruses from AID-KO BALB/cJ mammary tumors by Sanger sequencing.

Mutation	MMTV-WT Mutation Frequency[1]	MMTV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	0.80	1.65	2.1
A to G	0.56	0.75	1.3
C to T	**0.12**	**0.10**	**0.8**
T to C	0.64	1.28	2.0
Transition	2.12	3.78	1.8
Transversion	0.20	0.25	1.3
WRC	**0.08**	**0.08**	**1.0**
SYC	0.08	0.72	9.0
TYC	0.64	0.28	0.4
ATC	0.12	0.32	2.7

[1] Mutations/number of clones. Based on Sanger sequencing of 1,100 bp of the plus-strand of the proviral *env* gene clones obtained from five independent BALB/cJ AID-KO tumors induced by MMTV-WT (n=25) or MMTV-SD (n=40). The MMTV-SD clones analyzed retained the original SD mutation.

I also examined the number of proviral clones with reversion of the SD-site mutation as well as the distribution of mutations in different clones. Interestingly, 62.5% of the clones from MMTV-SD proviruses had reversion of the original SD mutation. This result is considerably higher than the percentage (22%) of clones obtained from MMTV-SD proviruses from mammary tumors appearing in wild-type BALB/cJ mice. One interpretation is that MMTV-SD replicates and recombines with endogenous *Mtv*s in a cell type enriched in AID-KO mice. The distribution of mutations/clone was not significantly different between the mutant and recombined MMTV-SD proviruses in WRC, TYC and ATC motifs (Figs. 3.17A, C and D). In contrast to results from MMTV-infected BALB/cJ wild-type mice, the number of mutations/clone in the SYC motif showed an enrichment in

MMTV-SD recombined proviruses (Fig. 3.17B). This result suggests that activity within the SYC motif is not due to AID, unlike reported previously [203, 204].

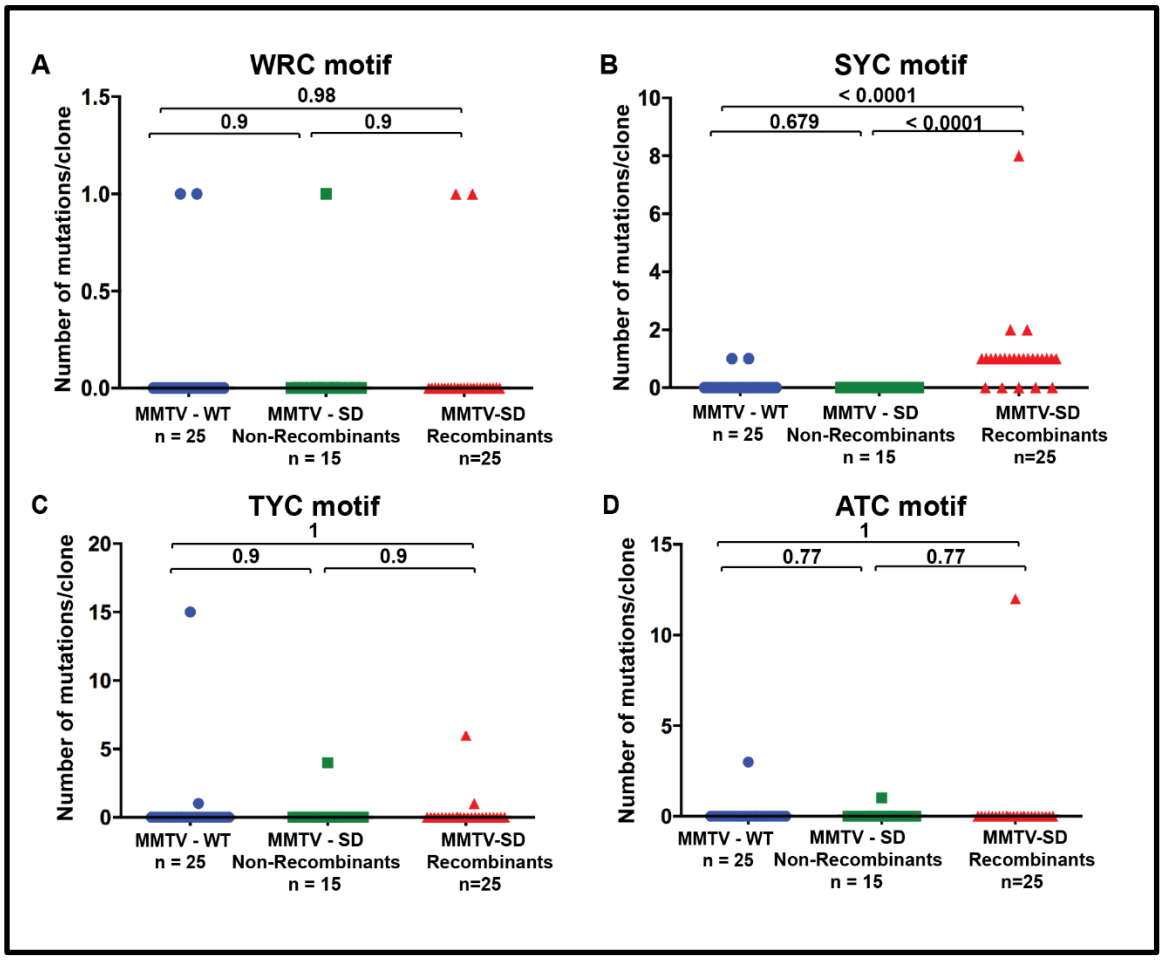

Figure 3.17: Mutational analysis of the proviral envelope gene from AID-KO BALB/cJ mammary tumors induced by MMTV-WT or MMTV-SD.

Number of C mutations within different motifs on either proviral strand is given for each clone. Clone numbers (n) are indicated. Sequences were obtained from independent clones from five tumors in different animals. (**A**) Number of mutations/clone in the WRC motif typical of AID mutation hotspots. (**B**) Number of mutations/clone in the SYC motif. (**C**) Number of mutations/clone in the TYC motif typical of mA3 mutation hotspots. (**D**) Number of mutations/clone in the ATC motif. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

Due to the lack of mutations in MMTV-SD recombinants recovered from tumors in AID-KO mice, correlation coefficients could not be calculated for WRC, TYC, and ATC motifs (Figs. 3.18 A, C and D). The correlation between SYC motif mutations and reversion of the SD site was however maintained in AID-KO mice (Fig. 3.18B). These data suggested that mutations in the SYC motif resulted from the activity of a deaminase distinct from AID.

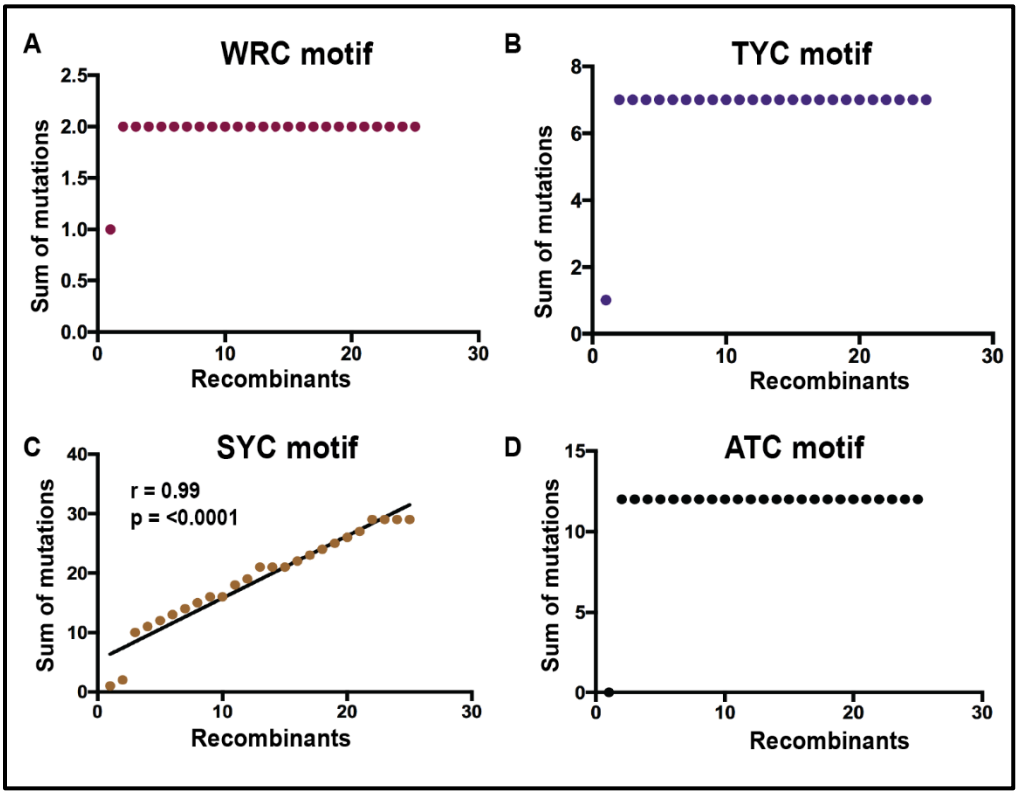

Figure 3.18: Correlation analysis of the number of recombinants and sum of mutations in different sequence motifs within MMTV-SD proviruses from BALB/cJ AID-KO mammary tumors.

(A, C and D) Spearman's correlation coefficient (r) could not be calculated between the number of recombinants and the number of mutations in WRC, TYC and ATC motifs, respectively, due to lack of mutations in SD recombinants. **(B)** Spearman's correlation coefficient (r) indicating high correlation between number of recombinants and number of mutations in SYC motif. Correlation coefficient p-values are shown on the graph.

Mutations within different sequence motifs were compared between MMTV-WT proviruses from mammary tumors in wild-type BALB/cJ mice and those in AID-KO mice. Direct analysis of the distribution of proviral mutations in the presence and absence of AID (Figs. 3.19 A-D) revealed that none were significantly different. In contrast, the distribution of mutations/clone in WRC and TYC motifs within MMTV-SD proviruses (MMTV-SD mutants and MMTV-SD recombinants combined) significantly declined in AID-KO mice compared to wild-type BALB/cJ. On the other hand, analysis of distribution of mutations/clone showed a significant increase within the SYC motif in MMTV-SD proviruses obtained from mammary tumors in AID-KO mice. Only mutations in the ATC motif showed a marginal, but insignificant, decline when comparing number of proviral MMTV-SD mutations/clone between BALB/cJ and AID-KO mice ($p = 0.08$) (Fig. 3.19D).

These results are consistent with the interpretation that the absence of Rem expression leads to increased mutations by AID and potentially other deaminases during MMTV transmission *in vivo*. Thus, the effect of loss of Rem on MMTV proviruses is similar to HIV-1 proviral hypermutation mediated by human APOBEC3 in the absence of Vif [97].

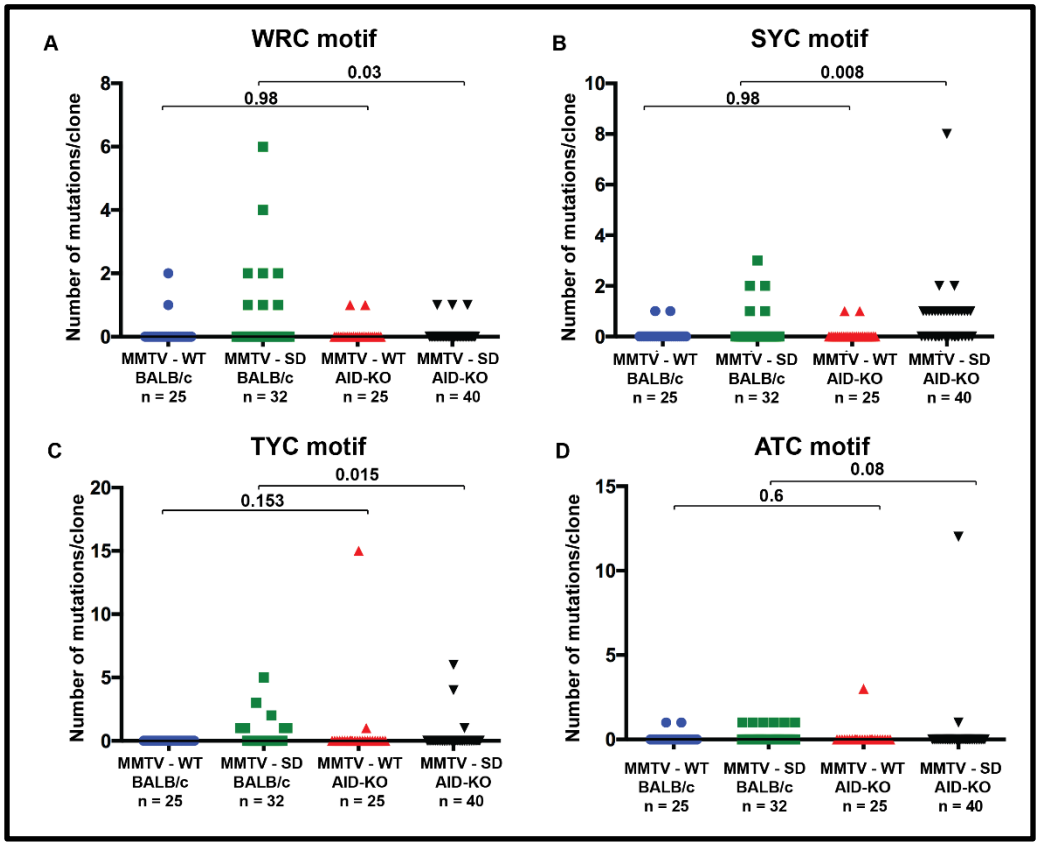

Figure 3.19: Comparisons of the distribution of mutations/clone within the proviral envelope gene between MMTV-WT and MMTV-SD-induced mammary tumors in either wild-type or AID-KO BALB/cJ mice.

Number of C mutations within different motifs on either proviral strand is given for each clone. Number of clones (n) is indicated. Sequences were obtained from independent clones from five tumors in different animals. **(A)** Number of mutations/clone in the WRC motif typical of AID mutation hotspots. **(B)** Number of mutations/clone in the SYC motif. **(C)** Number of mutations/clone in the TYC motif typical of mA3 mutation hotspots. **(D)** Number of mutations/clone in the ATC motif. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

3.2.5 AID is not incorporated into MMTV virions.

Specific human APOBEC3 proteins are packaged within budding HIV virions and restrict infection in target cells by deaminating negative-strand viral DNA during reverse transcription [97]. However, Vif expression in the producer cells targets APOBEC3 proteins for proteasomal degradation, thereby preventing their incorporation in the viral particle [153, 207]. Since murine AID belongs to the Apobec family, tha ability of mAID to restrict MMTV by the same mechanism was tested. The mA3 enzyme is packaged in MMTV particles isolated from virus-containing milk [179]. Thus, packaging of mA3 was used as a positive control for packaging of mAID.

To confirm packaging of mA3, 293T cells were co-transfected with CMV-driven MMTV expression plasmid (1 ug) and C-terminally HA-tagged mA3 plasmid (2 ug) using Lipofectamine 3000. Virus released from 293T cells was concentrated 100X with Takara RetroX concentrator. Western blotting was performed on concentrated virus before and after treatment with subtilisin to remove cellular proteins on the surface of virions. MMTV Gag-specific antibody detected capsid proteins in both cell extracts and in the concentrated supernatants (Fig. 3.20A, upper panel). As expected, an mA3-specific band was observed in cell extracts and in supernatants (Fig. 3.20A, lower panel). The presence of mA3 remained detectable after subtilisin treatment, confirming that mA3 is present within viral cores [179].

Similar experiments were performed to determine whether AID was packaged in MMTV particles by co-transfection of CMV-MMTV and murine AID (mAID) into 293T cells. Enrichment of the cleaved Gag protein was easily observed in cell supernatants

containing virion particles compared to extracts (Fig. 3.20B, upper panel). Western blotting with AID-specific antibody detected mAID in cell extracts, but not in cell supernatants (Fig. 3.20B, lower panel). These experiments suggest that mAID is not incorporated into wild-type MMTV virions.

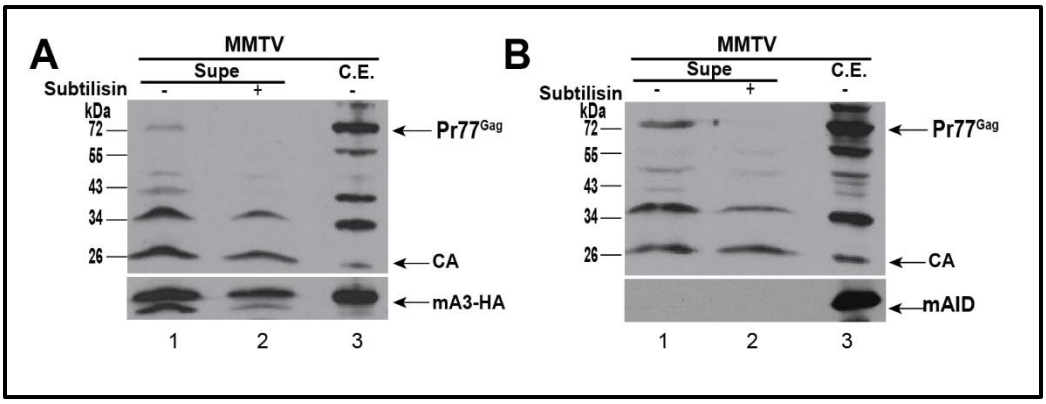

Figure 3.20: mAID is not incorporated into MMTV virions**.**

(**A**) mA3-HA is detectable in MMTV virions. Cells (293T) were co-transfected with MMTV and mA3-HA expressing plasmids. Western blots for both cell extracts (C.E.) and the 100-fold concentrated cell supernatant (-/+ subtilisin) are shown. Blots were incubated with either MMTV CA-specific (upper panel) or HA-specific (lower panel) antibody. (**B**) AID is not detectable in MMTV virions. Cells (293T) were co-transfected with MMTV and AID expressing plasmids. Western blots for both cell extracts (C.E.) and the 100-fold concentrated cell supernatant (-/+ subtilisin) are shown. Blots were incubated with either MMTV CA-specific (upper panel) or AID-specific (lower panel) antibody.

3.2.6 Rem causes proteasomal degradation of mAID.

HIV Vif is known to act as an adapter for a Cullin E3 ligase, resulting in the polyubiquitination and proteasomal degradation of human APOBEC3F and 3G [97]. Since AID was not detectably packaged within MMTV, Rem expression may affect the intracellular levels of mAID. Human Jurkat T cells, which lack endogenous *Mtv*s that may express Rem, were co-transfected with C-terminally GFP-tagged mAID expression plasmid and either wild-type or Rem-null virus expression plasmid. Western blotting revealed that Rem expression led to drastically reduced mAID-GFP levels (Fig. 3.21A, compare lanes 1 and 3). In contrast, in the absence of Rem expression mAID-GFP levels were maintained (Fig. 3.21A, compare lanes 1 and 5).

To confirm that Rem expression was responsible for decreased mAID levels, HEK293 cells were co-transfected with mAID-GFP and either N-terminally GFP-tagged Rem or untagged Rem plasmid. The presence of tagged or untagged Rem proteins lowered expression of mAID (Fig. 3.21B; compare lanes 1 and 3 or 5), whereas mAID levels were rescued in the presence of the proteasomal inhibitor MG-132 (Fig. 3.21B, even lanes). As previously reported, GFP-Rem precursor levels were stabilized by MG-132 compared to the cleaved product GFP-SP due to precursor susceptibility to endoplasmic reticulum-associated degradation (ERAD) (compare lanes 5 and 6) [24]. These results are consistent with a Vif-like activity of Rem.

Since mutations in mA3-specific TYC motifs were also higher in proviruses from Rem-null virus-induced tumors (Tables 3.1 and 3.2, Figs. 3.7 and 3.13), HEK293 cells were co-transfected with Rem expression vectors in the presence of plasmids expressing

mAID-GFP or mA3 C-terminally tagged with HA (mA3-HA). As expected, higher Rem levels led to a greater reduction of mAID levels (Fig. 3.21C, compare lanes 1-3). In contrast, the same Rem concentrations had no effect on mA3 expression (Fig. 3.21C, compare lanes 4-6). Thus, these data indicate that mAID is targeted for proteasomal degradation in the presence of Rem, whereas mA3 is not.

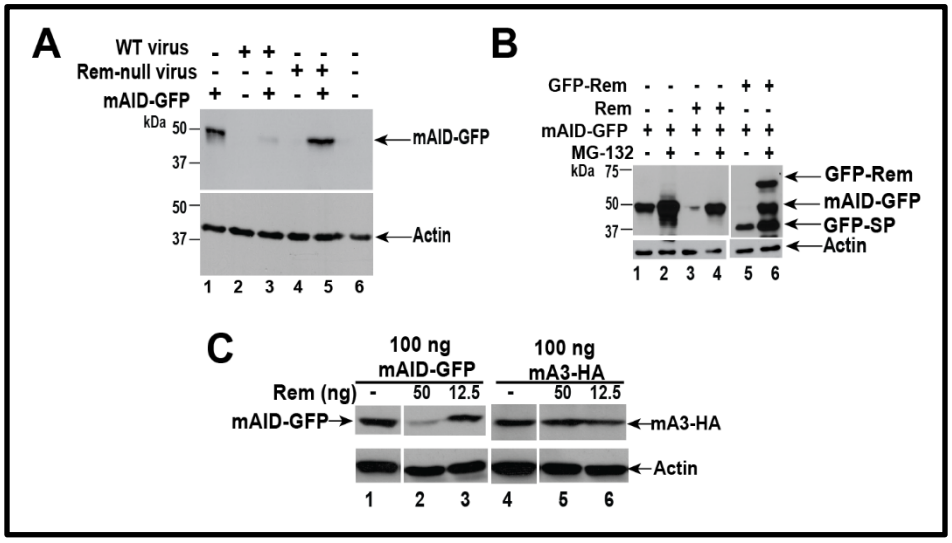

Figure 3.21: Rem antagonizes mAID.

(A) Co-transfection of Rem-expressing provirus leads to reduced mAID-GFP levels, whereas a Rem-null provirus does not. Either TBLV-WT or TBLV-SD expression plasmid was transfected into Jurkat cells in the presence or absence of murine AID tagged with GFP. Western blots were performed with GFP-specific (upper panel) or actin-specific (lower panel) antibody. **(B)** Rem expression leads to decreased mAID-GFP levels and is dependent on the proteasome. Cells (293) were co-transfected with expression plasmids for mAID-GFP and either GFP-tagged or untagged Rem. Samples in even lanes were prepared from cells treated with the proteasomal inhibitor MG-132. Western blots of cell extracts were incubated with GFP- or actin-specific antibody (upper and lower panels, respectively). **(C)** Rem expression does not affect mA3-HA levels. Cells (293) were transfected with the indicated amount of mAID-GFP expression vector or mA3-HA in the presence or absence of the indicated amount of untagged Rem expression plasmid. One blot (upper left) was incubated with GFP-specific antibody, whereas another blot (upper right) was incubated with HA-specific antibody. Actin-specific antibody (lower panels) was used to verify equal levels of protein loading. These results were obtained by Drs. Hyewon Byun and Frank Medina.

Taken together, these results are consistent with the interpretation that Rem protein is required for antagonizing the mutagenic effects of AID during MMTV replication prior to mammary gland transmission. Further mechanistic experiments are required to understand the stage of the MMTV life cycle for interactions with cellular Apobec enzymes. Moreover, the cellular compartment and host components needed for Rem interactions with murine cytidine deaminases remain to be determined.

Chapter 4: Apobec Enzyme Editing of Proviruses in TBLV-Induced T-Cell Tumors from C57BL/6 mice

4.1 RATIONALE

Studies in BALB/cJ mice showed that Rem antagonizes AID-associated proviral hypermutation, and also potentially the activity of other members of the Apobec family. Proviruses obtained from MMTV-SD (Rem-null)-injected BALB/cJ mice had increased mutations compared to MMTV-WT injected mice (Table 3.1 and Fig. 3.7). These mutations were observed in AID-specific WRC motifs as well as mA3-specific TYC motifs, and in two other motifs – ATC and SYC. In addition to Rem, the SD-site mutation also abolishes expression of *sag* transcript from the intragenic *env* but not the upstream LTR promoter [185]. Further, increased mutations were observed in TBLV-SD proviruses isolated from BALB/cJ T-cell tumors (Table 3.2 and Fig. 3.13). Since TBLV is a Sag-independent virus [70], increased proviral mutations were attributed to loss of Rem, not Sag expression. Also, transfection experiments showed that Rem co-expression leads to proteasomal degradation of AID (Fig. 3.21). MMTV requires replication in both B and T cells [30], which express AID and mA3. Therefore, viral genome hypermutation was analyzed after inoculation of MMTV-WT or MMTV-SD into AID-KO mice on the BALB/cJ background. Interestingly, the difference between MMTV-WT and MMTV-SD proviral mutations within both WRC and TYC motifs disappeared on the AID-KO background (Table 3.6, Fig. 3.17 and Fig. 3.19). Together, these results indicated that Rem

C-terminal sequences are critical for MMTV to antagonize Apobec family cytidine deaminases in BALB/cJ mice.

Mouse genetics have provided many important insights into the biology of viruses and the anti-viral immune response, including innate immunity mediated by Apobec family members [167-171]. Experiments using the RIII strain of MMTV to infect C57BL/6 (B6) mice were the first to show inhibition of retroviral replication and tumorigenesis by Apobec enzymes *in vivo* [179]. Interestingly, MMTV hypermutation was not observed in the presence or absence of mA3 [179]. Due to the availability of sequence data and ease of manipulating mouse genes, many of the engineered mouse strains are on the B6 background. Nevertheless, published data indicate that the levels of Apobec enzymes vary between BALB/c and B6 mice [180]. To further understand the involvement of cytidine deaminases in MMTV replication within various lymphocyte populations and in different mouse strains, I conducted additional experiments using the Sag-independent TBLV-WT virus and the Rem-null mutant (TBLV-SD) in B6 mice.

4.2 RESULTS

4.2.1 Tumors induced by Rem-null virus are accelerated in µMT B6 mice.

No difference in tumor incidence or latency was observed in BALB/cJ mice inoculated with TBLV-WT or Rem-null TBLV-SD, yet TBLV-SD-induced tumors had a decreased proviral load and increased mutations in WRC (AID-specific), TYC (mA3-specific), ATC and SYC motifs (Figs. 3.11 and 3.13). I sought to confirm this mutational phenotype in B6 mice, so that further studies could be performed in different mApobec

knockout strains on the B6 background to delineate the cell types and deaminases involved in mutating the viral genome in the absence of Rem. B6 mice lack the MHC class II I-E gene required for efficient C3H MMTV–Sag presentation and are thus less susceptible to MMTV infection [215]. Hence, TBLV, a Sag-independent variant of MMTV, which causes T-cell lymphomas in many mouse strains [51, 54, 55], was used for this study.

Wild-type B6 mice were injected intraperitoneally with 2 x 10^7 Jurkat cells expressing TBLV-WT or TBLV-SD (Rem-null) virus. Inoculation of TBLV-SD gave a slightly increased latency compared to TBLV-WT injection in B6 mice, which did not reach statistical significance (Fig. 4.1). Unlike BALB/cJ, in which 50% mice developed tumors (Fig. 3.11), almost all B6 mice developed tumors at the same dose of injection. This result suggests that B6 mice are more susceptible to TBLV infection. To confirm this observation, B6 mice were inoculated with $5x10^5$, $4x10^6$, and $2x10^7$ Jurkat cells producing TBLV-WT or TBLV-SD. As anticipated, the dose for B6 mice, which most closely approximated the tumor incidence in BALB/c animals, was $5x10^5$ TBLV-infected cells. Thus, the comparable dose for B6 mice was ~40 times lower than that for BALB/c (Table 4.1), indicating the high susceptibility of B6 mice to TBLV infection.

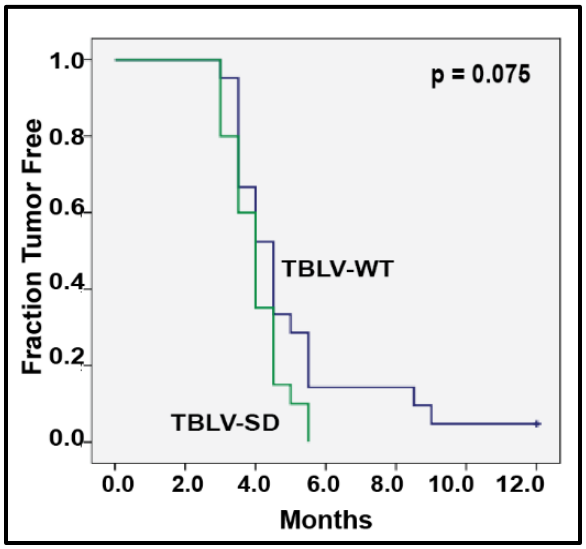

Figure 4.1: Kaplan-Meier plots for T-cell lymphoma development in B6 mice injected with TBLV-WT (blue line) or TBLV-SD (green line).

B6 mice develop T-cell tumors with a similar incidence and latency after infection with the Rem-null (TBLV-SD) virus (the p-value was marginal, but did not reach a statistically significant value of 0.05).

Table 4.1: T-cell lymphoma incidence in wild-type B6 mice at varying doses of virus injection.

Virus	Number of cells injected[1]	Tumor Incidence	Average Latency
TBLV-WT	$2x10^7$	100%	4 months
TBLV-SD	$2x10^7$	100%	4 months
TBLV-WT	$4x10^6$	75%	4 months
TBLV-SD	$4x10^6$	90%	3.3 months
TBLV-WT	$5x10^5$	50%	5 months
TBLV-SD	$5x10^5$	41%	5.5 months

[1] Mice were injected intraperitoneally with Jurkat cells stably expressing TBLV-WT or TBLV-SD.

To determine proviral hypermutation by Apobec enzymes, T-cell lymphomas were harvested from three independent TBLV-WT-induced tumors and an equal number of TBLV-SD-induced tumors (both groups injected with $2x10^7$ virus-producing cells). DNA extracted from these tumors was amplified using TBLV-specific primers and an *env*-LTR-specific PCR product was cloned. Individual clones were analyzed by Sanger sequencing. Sequencing results for the proviral plus strand are shown (Table 4.2). Contrary to results from TBLV infection in BALB/cJ mice (Table 3.2), differences between wild-type and Rem-null TBLV genome hypermutation generally were not observed in B6 mice, perhaps due to a dose of virus yielding a 100% tumor incidence. Nevertheless, similar to BALB/c mice, the frequency of ATC was increased for TBLV-SD proviruses (5-fold for B6 mice versus 2.6 fold for BALB/c mice). Mutational analysis of the viral genome will be repeated in this strain using a viral dose that results in a tumor incidence comparable to that observed in BALB/cJ mice (Fig. 3.11A).

Table 4.2: Mutation frequency in TBLV-WT and TBLV-SD proviruses from B6 thymic tumors by Sanger sequencing.

Mutation	TBLV-WT Mutation Frequency[1]	TBLV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	3.06	4.34	1.4
A to G	0.96	1.79	1.9
C to T	0.72	1.04	1.4
T to C	1.10	1.32	1.2
Total Transition	5.84	8.49	1.5
WRC	0.42	0.64	1.5
SYC	0.36	0.53	1.5
TYC	2.26	2.89	1.3
ATC	0.22	1.10	5.0

[1] Mutations/number of clones. Based on Sanger sequencing of 1,100 bp of the plus strand of the proviral *env* gene clones obtained from three independent B6 tumors induced by TBLV-WT (n=50) or TBLV-SD (n=53).

My previous data indicate that Rem-null MMTV and TBLV proviruses have increased WRC-motif mutations typical of AID relative to their respective wild-type proviruses in tumors from BALB/cJ mice (Tables 3.1 and 3.2). To determine if TBLV, like MMTV, requires replication in mature B cells where AID levels are reportedly the highest [216], C57BL/6 μMT (immunoglobulin heavy chain-knockout) mice and wild-type B6 mice were infected with TBLV-WT or TBLV-SD. This strain was developed by introducing a neomycin resistance cassette to disrupt one of the membrane exons of the gene encoding immunoglobulin heavy chain of the class IgM [220]. No expression of membrane-bound IgM is detectable together with a lack of peripheral B cells, although some B cells may be produced using a heavy-chain constant region gene other than IgM [221]. TBLV-WT gave a similar tumor incidence and latency as that observed for wild-type B6 mice (Fig. 4.2). Unexpectedly, lack of mature B cells led to a statistically

significant acceleration of T-cell tumors induced by TBLV-SD virus (Fig. 4.3). These results indicate that TBLV does not require replication in mature B cells for tumor induction in T cells.

To directly test the effect of AID on tumor development by TBLV in B6 mice, C57BL/6 AID-KO mice were also injected with $2x10^7$ TBLV-WT or SD-expressing Jurkat cells. No difference in latency or incidence was observed in TBLV-WT infected mice (Fig. 4.2). Unlike μMT mice, TBLV-SD did not display a significantly different tumor incidence or latency in AID-KO compared to the wild-type mice (Fig. 4.3). Lower doses of TBLV-WT and SD are also being tested to more closely mimic the susceptibility of B6 mice to this virus relative to BALB/c mice. Nevertheless, TBLV-SD-induced T-cell tumors developed with a shorter latency in μMT mice. This acceleration of T-cell lymphomas was observed in TBLV-SD infected μMT, and not AID-KO mice. Previous studies that reported mA3 as a restriction factor for MMTV in B6 mice showed that wild-type virus-induced mammary carcinomas were accelerated in mA3-deficient mice [179]. If a mature B-cell specific cytidine deaminase restricts TBLV, wild-type virus-induced tumors are also expected to accelerate in μMT mice, unlike the result obtained (Fig. 4.2). Thus, the existence of a restriction factor that antagonizes only Rem-null TBLV, a Sag-independent virus, in μMT mice is not in agreement with the predicted effects of known MMTV restriction factors.

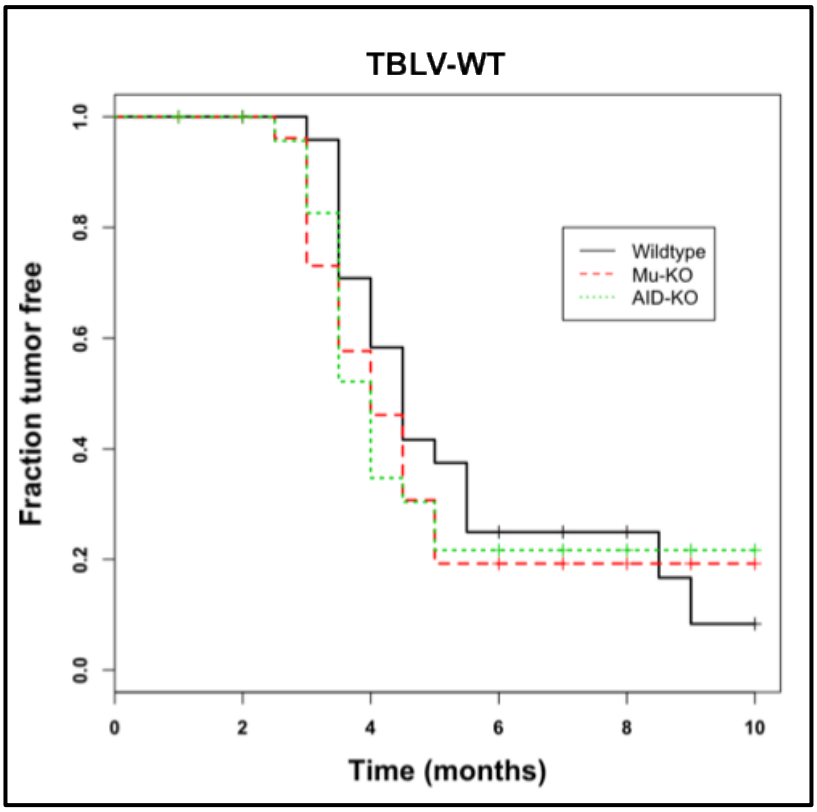

Figure 4.2: Kaplan-Meier survival plots for TBLV-WT-infected wild-type (black), AID-KO (green) and µMT (red) mice on B6 background.

The incidence and latency of T-cell lymphomas showed no significant difference between the three B6 strains. All p values were > 0.05, as tested by Mantel-Cox Log-rank test.

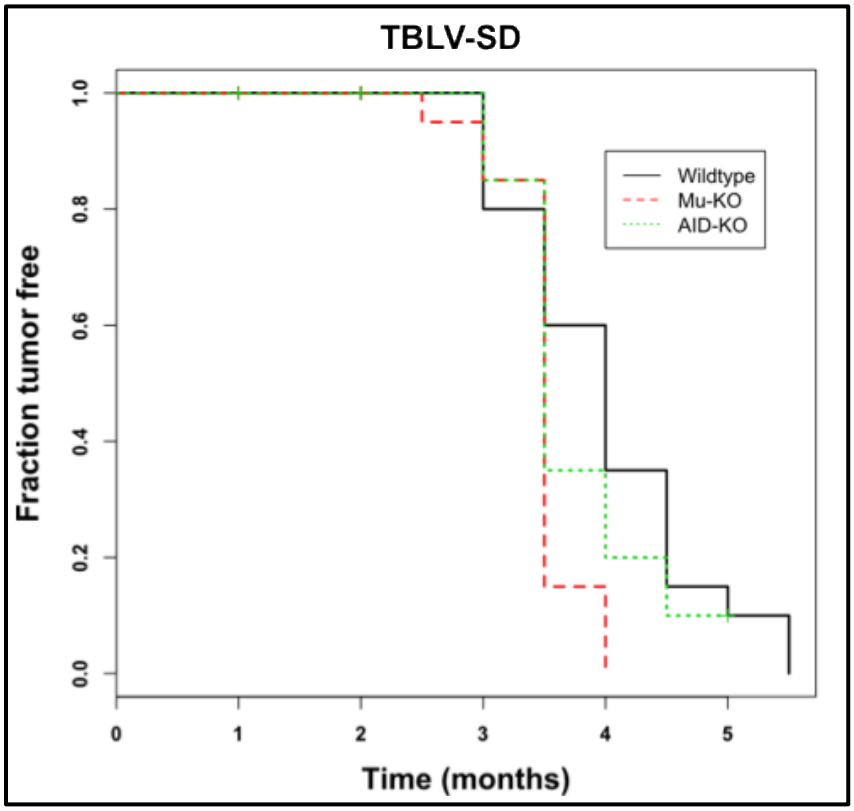

Figure 4.3: Kaplan-Meier survival plots for TBLV-SD-infected wild-type (black), AID-KO (green) and µMT (red) mice on B6 background.

The incidence and latency of TBLV-SD-induced T-cell lymphomas were significantly different only between wild-type and µMT B6 mice ($p = 0.006$). All other p values were > 0.05, as tested by the Mantel-Cox Log-rank test.

Studies have shown that mature B cells are required for efficient MMTV transmission after milk-borne infection [193]. However, development of T-cell tumors in µMT mice suggests that the life cycle of the T-cell variant TBLV does not require virus trafficking through mature B cells. Acceleration of tumor development in TBLV-SD-

infected µMT mice is unlikely due to lack of a good antibody response since this phenotype was not observed in TBLV-WT-infected µMT mice.

I also considered whether differential selection for *cis*-acting elements known to affect TBLV disease induction [64, 65] could accelerate TBLV-SD induction of tumors in µMT mice after Apobec-mediated mutagenesis. Specifically, TBLV proviruses have a triplicated T-cell enhancer element within the LTR that increases viral transcription in T cells [70]. A previous study has shown that changes within this region are selected during tumor passage, with four and one-repeat elements being transcriptionally less active [66]. As expected, PCR performed with LTR-specific primers showed a predominance of 3-repeat enhancers typical of the input virus. Further, no change in the number of enhancer repeats within TBLV-SD proviruses relative to TBLV-WT proviruses was observed in tumors from µMT mice that would confer an obvious selective advantage (Fig. 4.4).

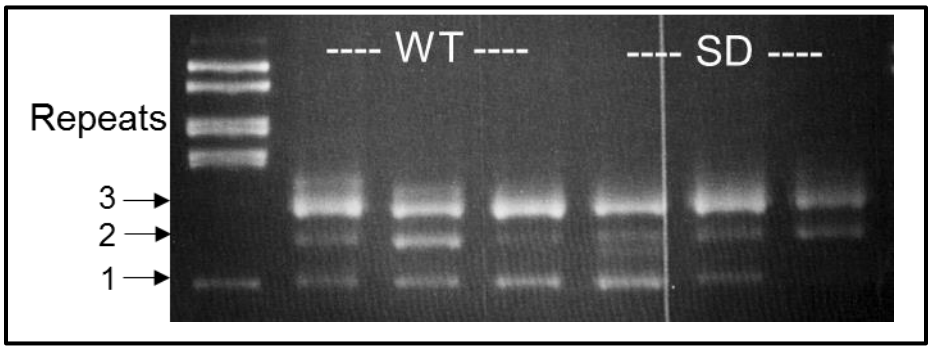

Figure 4.4: PCR analysis of LTR enhancer repeats from µMT B6 T-cell lymphomas induced by TBLV-WT or TBLV-SD.

PCR was performed on DNA extracted from three independent T-cell tumors using LTR 408+ and LTR786- primers. The numbers of repeats are indicated by arrows.

Sequencing results for the proviral plus strand from AID-KO and μMT mice showed that mutation frequency in the TYC motif was highest in all B6 strains (Tables 4.2, 4.3 and 4.4). Mutations within TYC motif showed an increase in the absence of Rem in μMT but not AID-KO mice. This suggests that absence of AID affects mutations within TYC motif, similar to the result in MMTV-infected BALB/c AID-KO mice (Table 3.6). Moreover, these results imply that the cytidine deaminase or cell type responsible for TYC motif mutations is enriched in μMT mice. Since the mutation frequency in WRC motif was higher in the absence of Rem in AID-KO mice, perhaps an enzyme different from AID can target this motif. In addition, mutations within ATC motif did not show an increase in the absence of Rem only in μMT mice. This suggests that the enzyme catalyzing ATC motif mutations is expressed in mature B cells and is antagonized by Rem.

Table 4.3: Mutation frequency in TBLV-WT and TBLV-SD proviruses from B6 AID-KO thymic tumors by Sanger sequencing.

Mutation	TBLV-WT Mutation Frequency[1]	TBLV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	2.38	2.09	0.9
A to G	0.58	0.6	1.0
C to T	0.22	0.35	1.6
T to C	0.78	0.65	0.8
Total Transition	3.96	3.69	1.0
WRC	0.16	0.60	3.8
SYC	0.54	0.56	1.0
TYC	1.58	1.42	0.9
ATC	0.16	0.55	3.4

[1] Mutations/number of clones. Based on Sanger sequencing of 1,100 bp of the plus strand of the proviral *env* gene clones obtained from three independent B6 AID-KO tumors induced by TBLV-WT (n=50) or TBLV-SD (n=55).

Table 4.4: Mutation frequency in TBLV-WT and TBLV-SD proviruses from B6 µMT thymic tumors by Sanger sequencing.

Mutation	TBLV-WT Mutation Frequency[1]	TBLV-SD Mutation Frequency[1]	Fold Increase (SD/WT)
G to A	3.74	7.59	2.0
A to G	0.74	0.57	0.8
C to T	0.54	0.43	0.8
T to C	0.98	0.45	0.5
Total Transition	6.00	9.04	1.5
WRC	0.72	1.02	1.4
SYC	0.52	0.69	1.3
TYC	3.10	5.61	1.8
ATC	0.40	0.41	1.0

[1] Mutations/number of clones. Based on Sanger sequencing of 1,100 bp of the plus strand of the proviral *env* gene clones obtained from three independent B6 µMT tumors induced by TBLV-WT (n=50) or TBLV-SD (n=51).

4.2.2 Rem-null virus-induced tumors have increased TYC consensus site mutations in µMT mice.

To address the increase in susceptibility of µMT mice to Rem-null virus, the mutational profile of the proviral genome was compared between tumors induced in wild-type and µMT mice. Although AID is known to be responsible for antibody affinity maturation in germinal center B cells, AID catalyzed SHM and CSR have been demonstrated in immature B cells as well [216]. AID expression has also been shown to be important for purging autoreactive B-cells from the immune repertoire [217]. TBLV-WT proviruses showed no significant difference in the distribution of WRC, SYC or ATC motif mutations between wild-type and µMT mice. TYC motif mutations typical of mA3 were slightly higher in the absence of mature B-cells (Fig. 4.5) ($p = 0.04$), perhaps indicative of TBLV infection of a different cell type.

The mutational profile of TBLV-WT proviruses from tumors induced in AID-KO mice on the B6 background were also compared to proviruses from T-cell tumors in wild-

type B6 mice. A significant decrease was observed in the WRC motif mutations by examination of the distribution of mutations/clone. Unlike the µMT mice, the distribution of TYC motif mutations/clone in TBLV-WT proviruses from wild-type and AID-KO tumors was unchanged (Fig. 4.6)

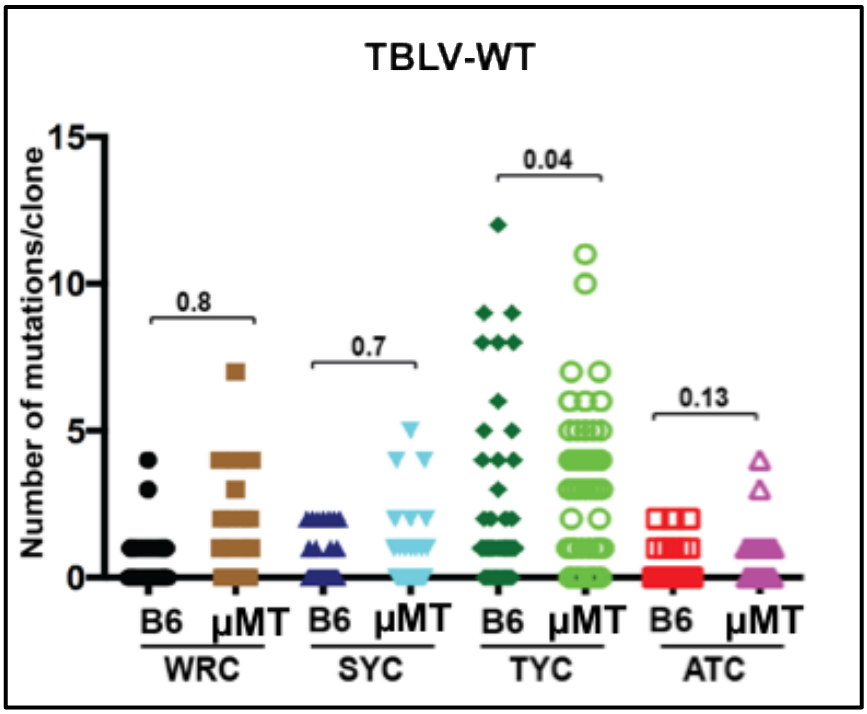

Figure 4.5: Comparison of the distribution of mutations within the proviral envelope gene from wild-type and µMT thymic tumors induced by TBLV-WT on the B6 background.

Independent cloned sequences were obtained from three tumors in different animals. Number of C mutations within different motifs on either proviral strand is given for each clone. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

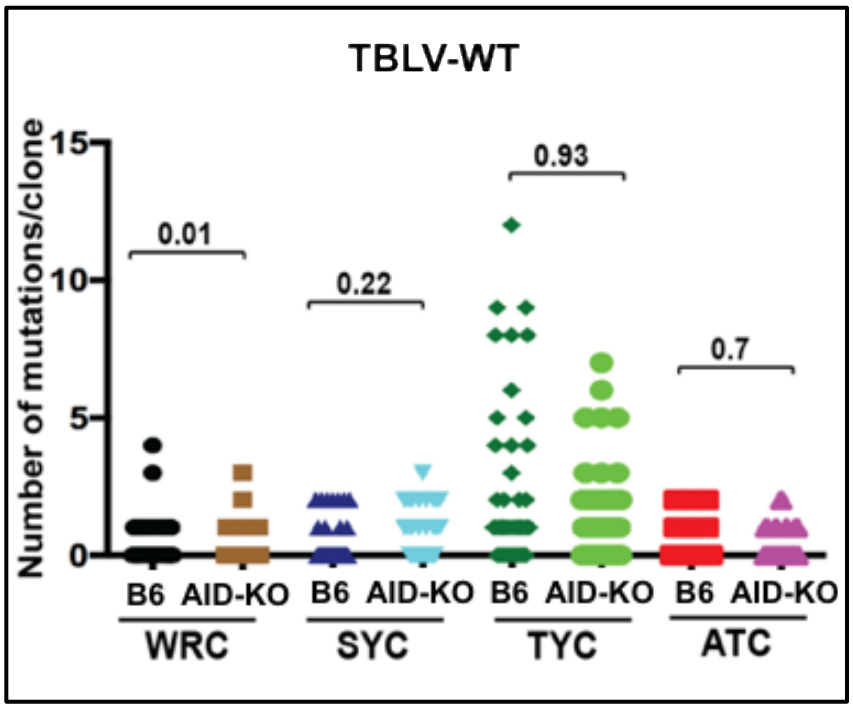

Figure 4.6: Comparison of the distribution of mutations within the proviral envelope gene from wild-type and AID-KO thymic tumors induced by TBLV-WT on the B6 background. Independent cloned sequences were obtained from three tumors in different animals. Number of C mutations within different motifs on either proviral strand is given for each clone. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

Next, mutations in all motifs were compared between TBLV-SD proviruses isolated from the three strains of mice. Distribution analysis revealed that mutations/clone in the ATC motif were reduced in µMT mice, which could be due to the absence of a cytidine deaminase specific for mature B cells. However, mutations/clone in the TYC motif were highly enriched in µMT mice compared to those from wild-type B6 mice, presumably due to increased transmission of virus through T-cells in mature B-cell deficient mice (Fig.

4.7). Taken together, these results fail to explain how tumors are accelerated in Rem-null infected µMT mice.

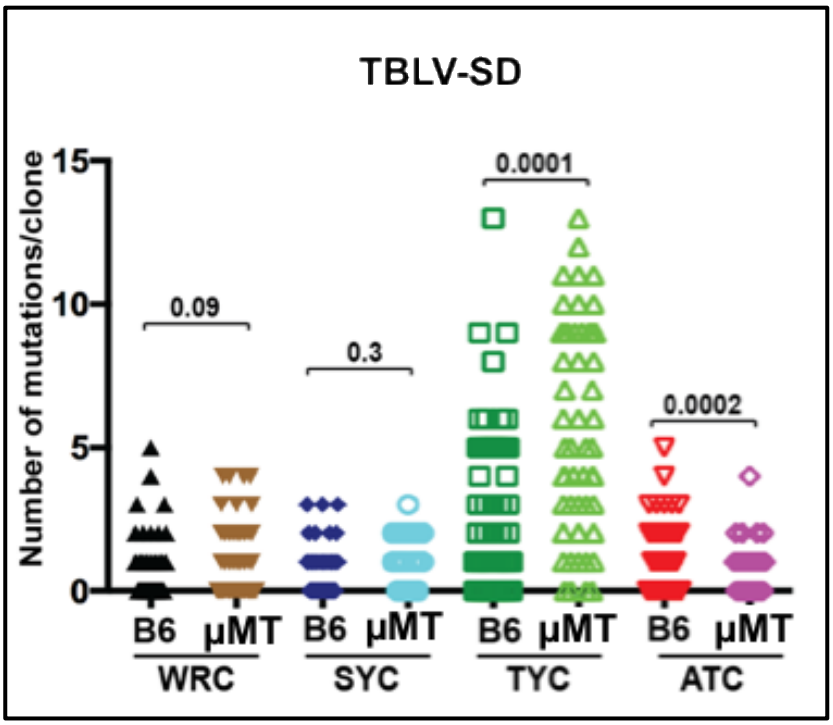

Figure 4.7: Comparison of the distribution of mutations within the proviral envelope gene from wild-type and µMT thymic tumors induced by TBLV-SD on the B6 background.

Independent cloned sequences were obtained from three tumors in different animals. Number of C mutations within different motifs on either proviral strand is given for each clone. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

TBLV-SD proviruses from T-cell lymphomas in AID-KO mice on the B6 background also were examined for different motifs targeted by cytidine deaminases. Distribution analysis revealed that both TYC and ATC motif mutations/clone were decreased significantly (Fig. 4.8). Unlike the mutational pattern for TBLV-WT, no significant difference was observed for the AID target motif, WRC. This mutational pattern

was compared to the analysis of MMTV-WT and MMTV-SD (Rem-null) proviruses in BALB/c AID-KO mice (Table 3.4 and Fig. 3.17). The absence of both AID and Rem in each situation reduced mutations in the WRC motif, but also unexpectedly reduced TYC and ATC motif mutations.

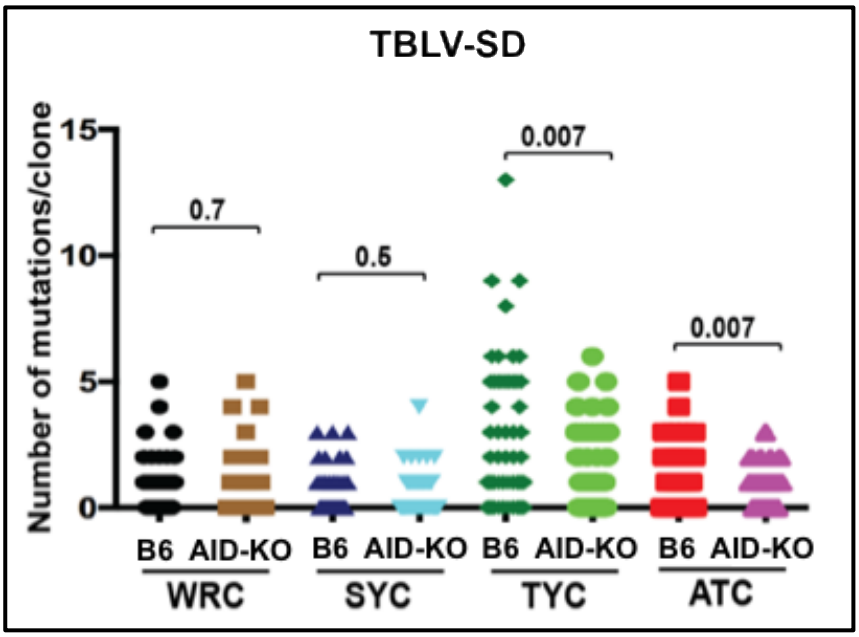

Figure 4.8: Comparison of the distribution of mutations within proviral envelope gene from wild-type and AID-KO thymic tumors induced by TBLV-SD on the B6 background.

Independent cloned sequences were obtained from three tumors in different animals. Number of C mutations within different motifs on either proviral strand is given for each clone. Statistical significance by non-parametric Mann-Whitney tests is indicated on the scatter plots.

4.2.3 Splenocytes from AID-GFP transgenic mice on B6 background can be stimulated *ex vivo* to induce and monitor AID expression.

Proviruses isolated from MMTV and TBLV-induced tumors harbored mutations in previously characterized as well as uncharacterized motifs. In both cases, viral genomes were analyzed for mutations in tumor-bearing cells. Induction of MMTV-induced mammary carcinomas likely occurs long after virus encounters restriction by deaminases during virus amplification in the lymphocyte population [30]. Thus, mechanistic understanding of the involvement of cytosine deaminase enzymes in retrovirus restriction can be better achieved with a tissue-culture system in which specific lymphocyte populations can be sorted at earlier time-points post-infection.

Rem-null proviruses showed mutations within motifs that are not specific for AID activity (non-WRC), suggesting that other deaminases can mutate the proviral genes (Table 3.1). However, mutations within the non-WRC motifs were reduced in AID-KO BALB/cJ mice (Table 3.6). This raises the possibility that AID acts on non-WRC motifs unlike previously published data for immunoglobulin genes [202]. Expression of AID has been reported outside of GC B cells [128, 216, 217]. Thus, using GFP as a marker will facilitate sorting for AID-expressing and non-expressing cells. C57BL/6 AID-GFP mice have been obtained from the Jackson Laboratory [218]. The GFP-sorted cells can be independently analyzed for mutations within all motifs to determine if non-WRC mutations are observed in MMTV proviruses only in cells not expressing AID.

To confirm that AID expression could be induced and monitored, splenocytes were obtained from an *Aicda-gfp* positive animal confirmed by genotyping of tail DNA as described [218]. Splenocytes were cultured in media containing lipopolysaccharide (LPS) and interleukin-4 (IL-4), and AID expression in cell lysates was confirmed at day 5 and day 10 by Western blotting (Fig. 4.9A). Incubation with AID-specific antibody yielded a

50 kDa band corresponding to AID-GFP in cells grown in stimulation media. A weaker 24 kDa band was observed with the expected size of endogenous mAID. These results suggested that the levels of AID-GFP are much higher than the endogenous AID protein, although it is not clear whether GFP tagging will affect hypermutation of MMTV proviruses.

Splenocytes from two independent AID-GFP positive mice were analyzed by flow cytometry using BD LSRII Fortessa Flow Cytometer to confirm GFP expression in induced splenocytes (Fig. 4.9B). An AID-GFP negative littermate was used as a control. Flow cytometry analysis suggested that there were two different populations of GFP-positive cells. Further characterization will be required to determine if both populations of cells are infected with MMTV and subjected to hypermutation.

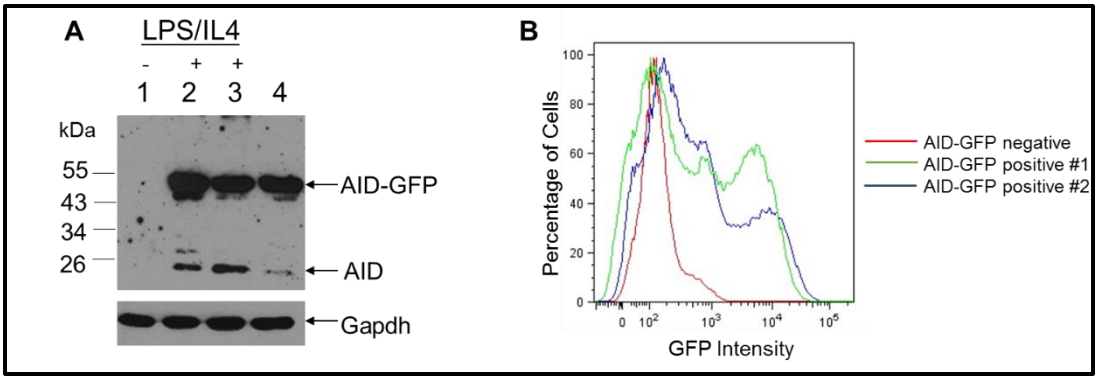

Figure 4.9: Induction of mAID in AID-GFP C57BL/6 transgenic mice.

(A) Splenocytes from AID-GFP positive mice were grown in regular media (lane 1) or media containing LPS and IL-4 (lanes 2 and 3) to induce AID. Lysates were analyzed by Western blotting with AID-specific antibody. A lysate from AID-GFP transfected 293T cells was used as a positive control (lane 4). **(B)** Splenocytes from AID-GFP negative (red) and positive littermates (green and blue) were grown in stimulant media for 4 days and analyzed by flow cytometry.

Chapter 5: Discussion and Perspective

5.1 MMTV REM ANTAGONIZES ACTIVATION-INDUCED CYTIDINE DEAMINASE (AID).

MMTV-Rem is a precursor that specifies an N-terminal Signal Peptide (SP), which is also generated by cleavage from envelope (Env) protein to serve a Rev-like function [24]. The role of the remaining Rem C-terminal sequences has not been described, but has been shown to be dispensable for virus replication in XC cells [185]. Therefore, I considered whether Rem C-terminal sequences serve an accessory function that facilitates MMTV replication *in vivo*.

Prior studies have revealed that the incidence of mammary tumors and viral loads were higher in mice deficient for mA3 [179]. However, no MMTV-encoded factor that antagonizes mA3 has been identified. Deaminases belonging to the Apobec family are expressed in B and T cells, both of which are required by MMTV for efficient transmission to the mammary gland [193, 25]. The potential role of Rem C-terminal sequences to counteract Apobec-mediated restriction of MMTV replication was assessed.

The Rem-null MMTV (MMTV-SD) was engineered as a splice donor mutant, which altered a single amino acid in envelope protein and eliminated production of spliced *rem* mRNA (Fig. 3.3B) as well as spliced *sag* mRNA from the intragenic *env* promoter [185]. Because the single amino acid change in Env did not affect viral RNA or protein production in rat XC cells [185] or virus release from human Jurkat cells (Fig. 3.10), the major effects of the splice donor mutation are likely due to failure to express either *sag* mRNA or *rem* mRNA or both. Therefore, two different viruses, the Sag-dependent MMTV

that causes mammary tumors [185] and the Sag-independent virus that induces T-cell lymphomas (TBLV) [70] were used to determine whether Rem is an Apobec antagonist.

Experiments conducted with wild-type and Rem-null MMTV demonstrated that Rem C-terminal sequences are needed for antagonism of the mutagenic effects of mAID. Multiple lines of evidence support this conclusion. First, tumors induced by wild-type MMTV had higher proviral loads than tumors induced by Rem-null virus (MMTV-SD) (Fig. 3.6B). Second, the number of C to T transitions and mutations within the AID-specific WRC motif were enriched in proviruses obtained from tumors induced by MMTV-SD compared to MMTV-WT in wild-type BALB/cJ mice (Table 3.1).Third, the difference in proviral load was abolished in tumors induced by MMTV in AID-KO BALB/cJ mice (Fig. 3.15B). Fourth, increased C to T transitions and WRC mutations were not observed in MMTV-SD proviruses from AID-KO BALB/cJ mice (Table 3.6 and Fig. 3.19A). Fifth, the Sag-independent variant TBLV [70] also showed lower proviral load and increased WRC-motif mutations in tumors induced by the cognate Rem-null virus (Fig. 3.11B and Table 3.2). Sixth, Rem co-expression led to proteasomal degradation of AID (Fig. 3.21). These results support the idea that Rem antagonizes AID during virus replication in lymphoid cells prior to mammary gland transmission.

Both tumor incidence and proviral loads were significantly reduced in Rem-null MMTV-injected BALB/cJ mice. On the other hand, BALB/cJ AID-KO mice lost the difference in proviral loads, but maintained the difference in tumor incidence. One possible explanation is that proviral loads are a Rem-dependent phenomenon since the absence of

Rem allows accumulation of mutagenized proviruses, which affects virus production. On the other hand, tumor incidence may be attributed to reduced Sag production from MMTV-SD [185] and decreased virus amplification in the lymphoid cell population. Lymphoma induction by the Sag-independent TBLV was unaffected by Rem expression relative to induction by TBLV-WT, but proviral loads were reduced in Rem-null TBLV-induced tumors. These data suggest that Rem expression increases the infectivity of viruses released from B and T cells prior to Sag-mediated transmission to the mammary gland.

The mutational phenotype of MMTV-SD proviruses in BALB/cJ mice also suggests the role of Rem C-terminal sequences in antagonizing mA3. In addition to the WRC motif, mutations within the mA3-specific TYC motif were also enriched in Rem-null virus induced tumors (Tables 3.1 and 3.2). Surprisingly, proviral mutation differences in both WRC and TYC motifs disappeared in tumors from AID-KO BALB/cJ mice (Table 3.6 and Fig. 3.19). Although TYC motif mutations are considered to be a result of mA3 activity, AID may cause these mutations as well. However, statistical analysis of distribution of mutations/clone in wild-type BALB/cJ mice showed that cytidine deamination within TYC motifs was elevated in tumor-derived MMTV-SD proviruses, regardless of SD-site recombination. On the other hand, cytidine deamination within proviral WRC motifs was selectively elevated only after SD-site recombination (Fig 3.7). Thus, a more likely interpretation is that TYC and WRC mutations occur in different cell types, and that AID ablation skews the abundance of cell types where MMTV replication occurs prior to infection of the mammary gland. Hence, the data also predict that AID has

direct or indirect effects on mA3-mediated mutagenesis of MMTV proviruses. Since both AID and hA3G independently form RNA-dependent high molecular mass (HMM) complexes in the cytosol [222, 223], it is possible that mAID and mA3 are part of the same complex. In that case, absence of mAID in AID-KO BALB/cJ mice will impact the expression or activity of mA3 and provide an explanation for the loss of mutations within the TYC motif.

Absence of Rem in MMTV-SD and TBLV-SD proviruses also yielded increased mutations in the ATC motifs (Tables 3.1 and 3.2). *In vitro* experiments have reported the ATC motif as an mA3 target specifically in BALB/cJ mice [180]. However, the TYC and ATC motif mutations did not always follow the same distribution of mutations/clone. For example, TYC motifs were significantly mutated between MMTV-WT and SD non-recombinants (proviruses than retained the SD-site mutation), whereas ATC motifs were not significantly altered when the same sets of clones were compared (Fig. 3.7). The same trend was observed when comparing the distribution of ATC motif mutations between TBLV-WT and SD-non recombinants (Fig. 3.13). These results suggest that TYC and ATC motif mutations may be due to different enzymes.

Proviruses from MMTV-SD-induced mammary carcinomas in BALB/cJ mice lacked significant enrichment in the SYC motif mutations compared to the wild-type proviruses (Fig. 3.7). On the other hand, proviruses from AID-KO BALB/cJ mice showed an increase in SYC-motif mutations (Fig. 3.17) in Rem-null tumors. Previous studies report sequence changes within the SYC motif as “cold-spot” mutations by AID [203, 204]. My

results imply that SYC-motif mutations on the MMTV provirus are due to the activity of another unidentified cytidine deaminase.

Interestingly, mutations in all motifs were more enriched in proviruses that had repaired the SD mutation. BALB/cJ mice have three endogenous *Mtv* copies, *Mtv6, 8* and *9* [32]. Only *Mtv8* and *Mtv9* are complete proviruses that are expressed in lymphocytes [27]. The commonly observed "reversion" of the SD mutation at all 6 mutant bases likely occurs in B and/or T cells [28, 185, 205, 208]. Importantly, occurrence of such an event indicates a strong selection for the restoration of the *env* gene SD site, which results in production of either *rem* mRNA from the LTR promoter or *sag* mRNA from the intragenic *env* promoter [185, 209]. Since reversion was observed in proviruses obtained from tumors induced by both Sag-dependent MMTV and Sag-independent TBLV, the data argue that the recombination event occurs to allow expression of Rem, not Sag. The frequency of SD-site reversion in TBLV-SD proviruses (9%) was lower than that in MMTV-SD proviruses (22%), likely due to the requirement for the original SD site exclusively for Rem production during of MMTV replication in B cells *in vivo*. The likelihood is low that reversion occurs to promote wild-type envelope protein production since proviral loads in MMTV-SD and WT-induced tumors in AID-KO mice were the same (Fig. 3.15B), indicating no replication effect of the single amino-acid Env mutation.

Rem production likely provides a selective advantage for virus replication in lymphocytes prior to Sag-mediated transmission to the mammary gland. A statistically significant correlation was found between the number of SD recombinants in both MMTV and TBLV proviruses and the number of mutations in WRC or TYC motifs (Fig. 3.14).

These data suggest that SD reversion by recombination with endogenous *Mtv*s and Apobec-induced mutations are occurring in cell types that are infected by both MMTV and TBLV. Moreover, MMTV recombinants in mammary tumor DNA had no stop codons in the *env* gene compared to TBLV recombinant proviruses in thymic tumors (compare Figs. 3.8 vs 3.12), consistent with the need to select infectious MMTVs during multiple cycles of replication in mammary tissue prior to tumor induction. Further experiments analyzing viral genome mutation and recombination within proviruses, especially those isolated at earlier time points prior to mammary tumor induction, are required to determine the lymphocyte sub-type where Apobec-mediated deamination is occurring.

Tissue culture experiments conducted with both tagged and untagged Rem expression plasmids showed drastic reduction in levels of AID in the presence of Rem. More importantly, these levels could be rescued by addition of the proteasomal inhibitor MG-132, strongly suggesting that Rem is a Vif-like factor that antagonizes the restriction factor AID (Fig. 3.21). Unlike AID, mA3 levels were not affected by Rem co-expression. Since TYC-motif mutations were reduced in AID-KO BALB/cJ mice, AID and mA3 may possibly be part of a larger Apobec complex. One possibility is that Rem-mediated degradation of AID yields an indirect effect on the stability of mA3. AID was not incorporated into MMTV virions, and increased viral genome mutation was not observed during *in-vitro* transfection experiments in AID-expressing Jurkat cells (Fig. 3.20 and Table 3.5). Hence, it is possible that upon infection, AID gains access to the MMTV genome following reverse-transcription. Nevertheless, these results represent the first AID-mediated antagonism of a retrovirus in a deamination-dependent manner.

In summary, these data strongly suggest that the retroviral protein Rem is an MMTV-encoded antagonist of AID in mice. Infected B and T cells act as a reservoir for MMTV until mice reach puberty, and the virus has access to actively dividing mammary epithelial cells. Lack of Rem results in increased mutations of the viral genome in lymphocytes. Proviruses with non-lethal mutations get selectively transmitted to the mammary gland, leading to mammary tumor induction (Fig. 5.1). AID is best known for its role in adaptive immunity in response to infectious agents [93, 210, 211], yet also has been shown to be important for restricting replication of herpesviruses in B cells [134]. Since AID overexpression has been associated with multiple human cancers [212], particularly B-cell lymphomas [213, 214], elucidation of the Rem-mediated mechanism targeting AID for degradation may be useful for both treatment of human tumors [212, 213] and infectious diseases.

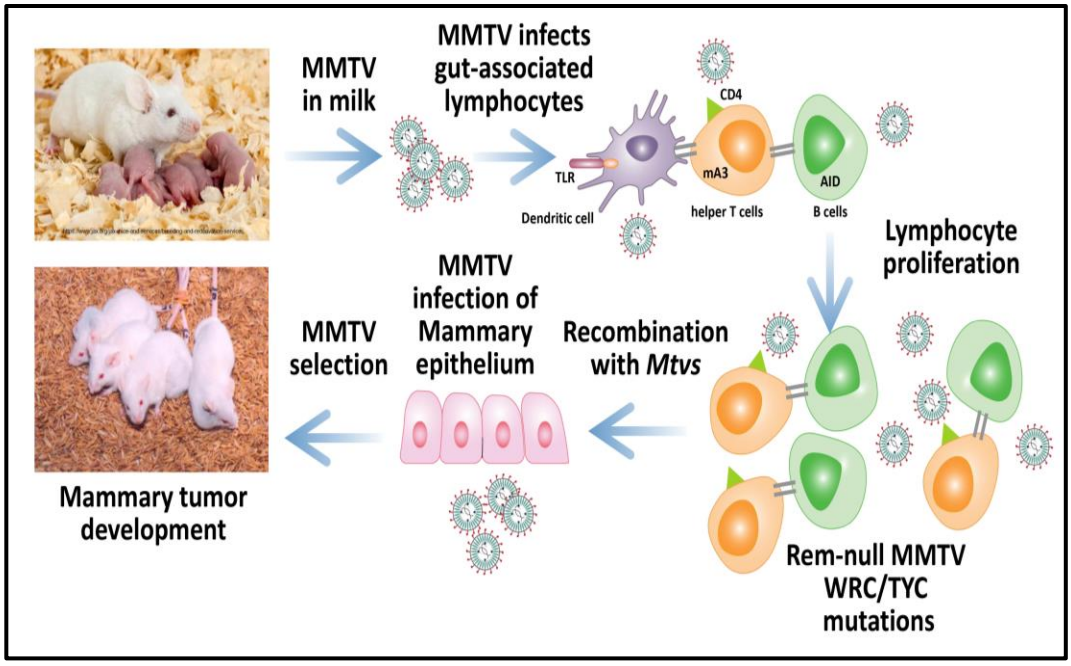

Figure 5.1: Proposed model for mApobec-mediated restriction of MMTV.

MMTV is transferred from infected mothers to pups during lactation. Virus infects dendritic and B cells, perhaps in gut-associated Peyer's patches. Superantigen expression on these antigen-presenting cells elicits a T-cell response that results in proliferation of infected lymphocytes. Rem-null viruses are subjected to increased mutations within WRC/TYC motifs by cytidine deaminases and undergo recombination with endogenous *Mtv*s. Viruses with non-lethal mutations are selected and transmitted to the mammary epithelial cells resulting in carcinomas by insertional mutagenesis.

5.2 MECHANISM OF REM FUNCTION

HIV Vif antagonizes proteins belonging to the human APOBEC family of cytidine deaminases by targeting them for proteasomal degradation [151-153]. This accessory protein protects the HIV genome from APOBEC-mediated mutagenesis in the target cells following infection. Tissue culture experiments demonstrated that Rem co-expression results in proteasomal degradation of mAID (Fig. 3.21). It is not clear whether the precursor Rem or cleaved Rem CT protein is responsible for AID antagonism. Uncleaved Rem is retrotranslocated to the cytoplasm, ubiquitylated and subjected to ERAD [24]. Upon cleavage, a small fraction of Rem-CT utilizes ERAD to retrotranslocate to the cytoplasm. Another portion of the protein leaves the ER but does not follow the typical secretory pathway [Wendy Xu, unpublished data]. The most stable levels of Rem-CT are seen in the ER. Thus, either full-length Rem or a portion of the cleaved Rem-CT may act as an adapter to target AID for degradation. Degradation of mA3 was not observed in the presence of Rem (Fig. 3.21). However, if AID and mA3 co-exist in a complex, Rem-mediated ubiquitylation of AID can result in destabilization of mA3 (Fig.5.2). Further experiments with Rem expression constructs that are resistant to cleavage will address the form of Rem protein required for AID degradation.

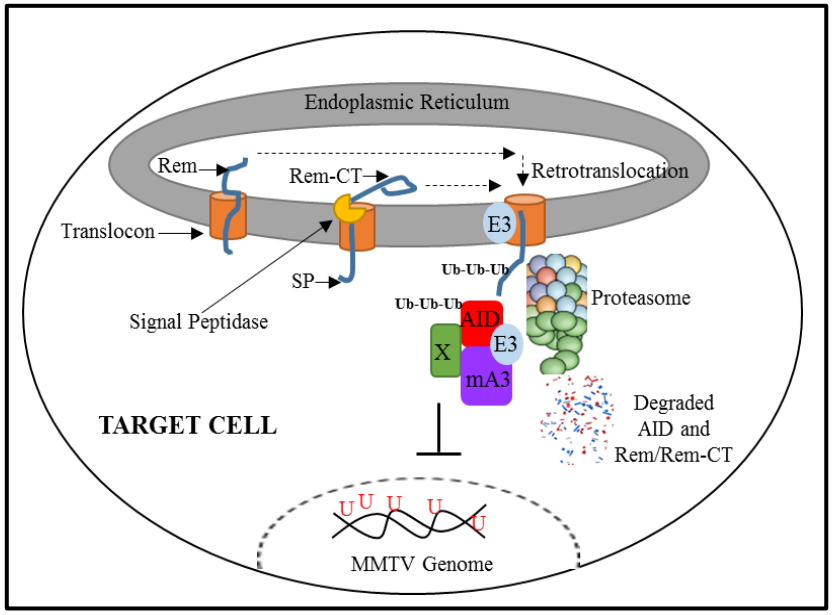

Figure 5.2: Proposed model for Rem-mediated antagonism of AID.

Rem is synthesized at the ER membrane and cleaved by signal peptidase into SP and Rem-CT [24]. The Rem-CT cleavage product is glycosylated in the ER lumen, where a fraction is retrotranslocated for degradation by cytosolic proteasomes. Rem precursor protein is also ubiquitylated and subjected to ERAD. Retrotranslocated Rem or Rem-CT acts as a Vif-like adaptor to enable addition of ubiquitin chains on AID in cytosolic complexes with mA3 and other Apobecs (X) causing proteasomal degradation of AID, thus inhibiting mutation of MMTV proviruses.

Human APOBEC proteins deaminate cytidines on the HIV genome by incorporating into budding HIV virions in the absence of Vif [135-140]. Studies have shown mA3 to be packaged into MMTV virions [179], but experiments conducted in tissue culture with wild-type MMTV expressing plasmid did not detect AID in the mature MMTV particles (Fig. 3.20). However, it remains unknown whether AID gains access to the budding viral particle in the absence of Rem expression. In this case, if AID is packaged within MMTV virions, Rem will be mechanistically analogous to HIV Vif protein. On the other hand, if AID is not incorporated within Rem-null MMTV, further experiments will be necessary to determine the stage of the virus life cycle when AID accesses the viral genome. Mutational analysis performed on proviruses from human Jurkat T cells stably expressing AID did not show increased viral genome mutation in the absence of Rem (Table 3.5). Since human Jurkat cells are not re-infected due to their failure to express the mouse TfR1 receptor for MMTV Env protein [64], absence of increased mutagenesis indicates that AID acts on MMTV in the target cell following reverse transcription, perhaps as part of the viral pre-integration complex (PIC). *In vitro* experiments will be conducted, both with inhibitors of viral reverse transcription and proviral integration, to determine the exact stage of the virus life cycle when AID targets the viral genome.

5.3 APOBEC ENZYME EDITING OF PROVIRUSES IN TBLV-INDUCED T-CELL TUMORS FROM C57BL/6 MICE

Since Rem-null proviruses recovered from BALB/cJ mice had mutations in non-WRC motifs (Tables 3.1, 3.2 and 3.6), experiments in mouse strains lacking mA3 or mA1 would shed light on the role of non-AID cytidine deaminases in restricting MMTV. Most such single or double-knockout strains are on a C57BL/6 (B6) background. Therefore, experiments were conducted in wild-type B6 mice to confirm the mutational phenotype observed in BALB/cJ mice. Because B6 mice are unable to efficiently present C3H MMTV Sag due to a mutation in MHC class II genes [215], I used the Sag-independent virus, TBLV [70], for these studies. Surprisingly, the TBLV dose that gave a 50% tumor incidence in BALB/cJ mice resulted in a nearly 100% incidence in B6 mice (Figs. 4.1 and 3.11A). The reason for this difference in susceptibility is not known, but these two strains have been reported to be immunologically distinct [219]. In response to pathogens, T cells from B6 mice preferentially produce Th1 cytokines with high interferon-gamma (IFNγ) and low interleukin (IL-4), whereas those from BALB/cJ produce Th2 cytokines with low IFNγ and high IL-4. Proviruses analyzed from B6 tumors inoculated at the high dose resulted in much smaller differences in the numbers of mutations/clone between TBLV-WT and TBLV-SD proviruses compared to BALB/c tumor-derived proviruses (Table 4.2). Additional studies in the B6 strain are needed at a viral dose that results in an incidence and latency more comparable to that observed in BALB/cJ mice. Such experiments may result in a mutational profile similar to that observed with TBLV infection of BALB/cJ mice (Table 3.2), where absence of Rem resulted in increased proviral gene mutations.

The μMT B6 mice lacking mature B cells were injected with TBLV-WT or SD at the same time as the wild-type B6 strain. Similar experiments were performed in AID-KO B6 mice. No significant difference in tumor incidence or latency was observed after

inoculation of either knockout mouse strain compared to wild-type B6 mice after inoculation with TBLV-WT (Fig. 4.2). However, loss of mature B cells, but not AID, coupled with Rem deficiency accelerated the latency of T-cell lymphomas induced by TBLV (Fig. 4.3). These results suggest that Rem inhibits TBLV replication or its transmission to T cells, presumably during interactions with mature B cells, but this inhibition is independent of AID expression.

Proviruses from μMT tumors had increased mutations/clone only in the mA3-specific TYC motif within TBLV-WT proviruses relative to those obtained from wild-type B6 mice (Fig. 4.5). The same comparisons of TBLV-SD proviruses showed a highly significant increase in the number of TYC motif mutations/clone, but also a significant decrease in ATC motif mutations/clone (Fig. 4.7). As expected, TBLV-WT infected mice showed significantly fewer WRC motif mutations/clone in AID-KO B6 tumors, but the numbers of SYC, TYC and ATC mutations/clone were unaffected compared to those in wild-type B6 tumors (Fig. 4.6). The same analysis of TBLV-SD proviruses revealed no change in WRC motif mutations/clone between wild-type and AID-KO B6 mice, yet both TYC and ATC motif mutations/clone were significantly decreased (Fig. 4.8). These results may be due to altered transmission of TBLV through T cells as well as infection of different populations of B cells.

Ex vivo experiments were conducted using splenocytes harvested from AID-GFP transgenic mice on B6 background. Splenocytes were be grown in tissue culture conditions for almost two weeks and stimulated with LPS and IL-4 to induce expression of GFP-tagged AID. This was confirmed by Western blotting and flow cytometry (Fig. 4.9). Data from BALB/cJ mice indicates that mApobec-mediated viral genome hypermutation is occurring in the lymphocytes, likely long before tumor development. Thus, splenocytes isolated from AID-GFP-expressing transgenic animals injected with MMTV can be used

to study phenotypic changes during early stages of virus infection, such as (i) monitoring for AID expression in cells infected with wild-type compared to Rem-null MMTV, and (ii) analyzing proviral genome mutations in cells infected with wild-type and Rem-null MMTV. Different lymphocyte sub-populations can be sorted using GFP and other cell-surface markers to determine whether mutations in certain motifs occur in specific cell types and whether these mutations lead to decreased virus infectivity.

In summary, my *in vivo* studies combined with tissue culture experiments suggest that Rem is a Vif-like antagonist of multiple Apobec family enzymes. Future experiments in AID-GFP transgenic mice as well as in different mApobec knockout strains will help to delineate the specific cell types and enzymes involved in virus restriction as well as their mechanism of action.

Appendix

APOBEC	apolipoprotein B mRNA editing enzyme
AID	activation-induced cytidine deaminase
APC	antigen presenting cell
ARM	arginine-rich motif
BIV	bovine immunodeficiency virus
CA	capsid
CIS	common integration sites
Crm1	chromosome region maintenance 1 protein
CsCl	cesium chloride
CSR	class switch recombination
CTE	constitutive transport element
Cux1	CCAAT displacement protein 1
Da	daltons
DC	dendritic cell
DEPC	diethyl pyrocarbonate
Dex	dexamethasone
DMEM	Dulbecco's modified eagles medium
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
DU	dUTPase
EDTA	ethylenediamine tetraacetic acid
EBV	epstein-barr virus
Env	envelope
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum-associated protein degradation
ERGIC	ER-golgi intermediate compartment
FBS	fetal bovine serum
FIV	feline immunodeficiency virus
Gag	group-specific antigen
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GC	germinal center
GR	glucocorticoid receptor
HBV	hepatitis B virus
HERV	human endogenous retrovirus
HIV	human immunodeficiency virus
HPV	human papillomavirus
HRE	hormone-responsive element
HTLV	human T-cell leukemia virus

IL	interleukin
IN	integrase
KSHV	kaposi's sarcoma herpes virus
LB	luria-bertani broth
LINE	long interspersed nuclear elements
LPS	lipopolysaccharide
LTR	long terminal repeat
LUC	luciferase
mA1	murine Apobec1
mA3	murine Apobec3
MA	matrix
MEC	mammary epithelial cells
MG	mammary gland
MGE	mammary gland enhancer
MHC	major histocompatibility complex
miR	micro RNA
MMTV	mouse mammary tumor virus
MuLV	murine leukemia virus
Mtvr2	mouse mammary tumor virus receptor homolog 2
NC	nucleocapsid
NES	nuclear export sequence
NLS	nuclear localization signal
NoLS	nucleolar localization signal
NRE	negative regulatory element
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
Pol	polymerase
PR	protease
Rem	regulator of export of MMTV mRNA
Rev	regulator of virion expression
Rfv3	recovery from friend virus 3
RmRE	Rem-responsive element
RNP	ribonucleoprotein
RRE	Rev response element
RNA	ribonucleic acid
RPMI	Roswell park memorial institute
RT	reverse transcriptase/ room temperature
RT-PCR	reverse transcription-polymerase chain reaction
Sag	superantigen
SDS	sodium dodecyl sulphate
SHM	somatic hypermutation

SINE	short interspersed nuclear elements
SIV	simian immunodeficiency virus
SP	signal peptide
SU	surface envelope protein
TBLV	Type-B leukemogenic virus
TCR	T cell receptor
TfR1	transferrin receptor 1
TM	transmembrane envelope protein

Bibliography

1. Jackson, R. B., & Little, C. C. (1933). The existence of non-chromosomal influence in the incidence of mammary tumors in mice. *Science*, *78*(2029), 465-466.
2. Bitner, J. J. (1936). Some possible effects of nursing on the mammary gland tumor incidence in mice. *Science*, *84*(2172), 1962.
3. Graff, S., Moore, D. H., Stanley, W. M., Randall, H. T., & Haagensen, C. D. (1949). Isolation of mouse mammary carcinoma virus. *Cancer*, *2*(5), 755-762.
4. Schlom, J., Spiegelman, S., & Moore, D. H. (1972). Reverse transcriptase and high molecular weight RNA in particles from mouse and human milk. *Journal of the National Cancer Institute*, *48*(4), 1197-1203.
5. Scolnick, E. M., Aaronson, S. A., & Todaro, G. J. (1970). DNA synthesis by RNA-containing tumor viruses. *Proceedings of the National Academy of Sciences*, *67*(2), 1034-1041.
6. Callahan, R., & Smith, G. H. (2008). The mouse as a model for mammary tumorigenesis: history and current aspects. *Journal of Mammary Gland Biology and Neoplasia*, *13*(3), 269-269.
7. Briggs, J. A., Watson, B. E., Gowen, B. E., & Fuller, S. D. (2004). Cryoelectron microscopy of mouse mammary tumor virus. *Journal of Virology*, *78*(5), 2606-2608.
8. Sarkar, N. H., & Moore, D. H. (1974). Surface structure of mouse mammary tumor virus. *Virology*, *61*(1), 38-55.
9. Mertz, J. A., Chadee, A. B., Byun, H., Russell, R., & Dudley, J. P. (2009). Mapping of the functional boundaries and secondary structure of the mouse mammary tumor virus Rem-responsive element. *Journal of Biological Chemistry*, jbc-M109.
10. Reuss, F. U., & Coffin, J. M. (1995). Stimulation of mouse mammary tumor virus superantigen expression by an intragenic enhancer. *Proceedings of the National Academy of Sciences*, *92*(20), 9293-9297.
11. Zhu, Q., Maitra, U., Johnston, D., Lozano, M., & Dudley, J. P. (2004). The homeodomain protein CDP regulates mammary-specific gene transcription and tumorigenesis. *Molecular and Cellular Biology*, *24*(11), 4810-4823.
12. Ross, S. R. (2008). MMTV infectious cycle and the contribution of virus-encoded proteins to transformation of mammary tissue. *Journal of Mammary Gland Biology and Neoplasia*, *13*(3), 299.
13. Hizi, A., Henderson, L. E., Copeland, T. D., Sowder, R. C., Hixson, C. V., & Oroszlan, S. (1987). Characterization of mouse mammary tumor virus gag-pro gene products and the ribosomal frameshift site by protein sequencing. *Proceedings of the National Academy of Sciences*, *84*(20), 7041-7045.

14. Jacks, T., Townsley, K., Varmus, H. E., & Majors, J. (1987). Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. *Proceedings of the National Academy of Sciences*, *84*(12), 4298-4302.
15. Sarkar, N. H., Whittington, E. S., Racevskis, J., & Marcus, S. L. (1978). Phosphoproteins of the murine mammary tumor virus. *Virology*, *91*(2), 407-422.
16. Payne, S. L., & Elder, J. H. (2001). The role of retroviral dUTPases in replication and virulence. *Current Protein and Peptide Science*, *2*(4), 381-388.
17. Ross, S. R., Schofield, J. J., Farr, C. J., & Bucan, M. (2002). Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *Proceedings of the National Academy of Sciences*, *99*(19), 12386-12390.
18. Choi, Y., Kappler, J. W., & Marrack, P. (1991). A superantigen encoded in the open reading frame of the 3′ long terminal repeat of mouse mammary tumour virus. *Nature*, *350*(6315), 203.
19. Marrack, P., Kushnir, E., & Kappler, J. (1991). A maternally inherited superantigen encoded by a mammary tumor virus. *Nature*, *349*(6309), 524.
20. Golovkina, T. V., Dudley, J. P., & Ross, S. R. (1998). B and T cells are required for mouse mammary tumor virus spread within the mammary gland. *The Journal of Immunology*, *161*(5), 2375-2382.
21. Ross, S. R. (2000). Using genetics to probe host-virus interactions; the mouse mammary tumor virus model. *Microbes and Infection*, *2*(10), 1215-1223.
22. Indik, S., Günzburg, W. H., Salmons, B., & Rouault, F. (2005). A novel mouse mammary tumor virus encoded protein with Rev-like properties. *Virology*, *337*(1), 1-6.
23. Mertz, J. A., Simper, M. S., Lozano, M. M., Payne, S. M., & Dudley, J. P. (2005). Mouse mammary tumor virus encodes a self-regulatory RNA export protein and is a complex retrovirus. *Journal of Virology*, *79*(23), 14737-14747.
24. Byun, H., Halani, N., Mertz, J. A., Ali, A. F., Lozano, M. M., & Dudley, J. P. (2010). Retroviral Rem protein requires processing by signal peptidase and retrotranslocation for nuclear function. *Proceedings of the National Academy of Sciences*, *107*(27), 12287-12292.
25. Golovkina, T. V., Chervonsky, A., Dudley, J. P., & Ross, S. R. (1992). Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. *Cell*, *69*(4), 637-645.
26. Baillie, G. J., van de Lagemaat, L. N., Baust, C., & Mager, D. L. (2004). Multiple groups of endogenous betaretroviruses in mice, rats, and other mammals. *Journal of Virology*, *78*(11), 5784-5798.
27. Kozak, C. E. A. L., Peters, G., Pauley, R., Morris, V., Michalides, R., Dudley, J., & Vaidya, A. (1987). A standardized nomenclature for endogenous mouse mammary tumor viruses. *Journal of Virology*, *61*(5), 1651.

28. Golovkina, T. V., Piazzon, I., Nepomnaschy, I., Buggiano, V., de Olano Vela, M., & Ross, S. R. (1997). Generation of a tumorigenic milk-borne mouse mammary tumor virus by recombination between endogenous and exogenous viruses. *Journal of Virology*, *71*(5), 3895-3903.
29. Holt, M., Shevach, E., & Punkosdy, G. (2013). Endogenous mouse mammary tumor viruses (mtv): new roles for an old virus in cancer, infection, and immunity. *Frontiers in Oncology*, *3*, 287.
30. Barnett, A., Mustafa, F., Wrona, T. J., Lozano, M., & Dudley, J. P. (1999). Expression of mouse mammary tumor virus superantigen mRNA in the thymus correlates with kinetics of self-reactive T-cell loss. *Journal of Virology*, *73*(8), 6634-6645.
31. Scherer, M. T., Ignatowicz, L., Pullen, A., Kappler, J., & Marrack, P. (1995). The use of mammary tumor virus (Mtv)-negative and single-Mtv mice to evaluate the effects of endogenous viral superantigens on the T cell repertoire. *Journal of Experimental Medicine*, *182*(5), 1493-1504.
32. Bhadra, S., Lozano, M. M., Payne, S. M., & Dudley, J. P. (2006). Endogenous MMTV proviruses induce susceptibility to both viral and bacterial pathogens. *PLoS Pathogens*, *2*(12), e128.
33. Finke, D., & Acha-Orbea, H. (2001). Differential migration of in vivo primed B and T lymphocytes to lymphoid and non-lymphoid organs. *European Journal of Immunology*, *31*(9), 2603-2611.
34. Karapetian, O., Shakhov, A. N., Kraehenbuhl, J. P., & Acha-Orbea, H. (1994). Retroviral infection of neonatal Peyer's patch lymphocytes: the mouse mammary tumor virus model. *Journal of Experimental Medicine*, *180*(4), 1511-1516.
35. Xu, L., Wrona, T. J., & Dudley, J. P. (1996). Exogenous mouse mammary tumor virus (MMTV) infection induces endogenous MMTV sag expression. *Virology*, *215*(2), 113-123.
36. Xu, L., Wrona, T. J., & Dudley, J. P. (1997). Strain-specific expression of spliced MMTV RNAs containing the superantigen gene. *Virology*, *236*(1), 54-65.
37. Ardavín, C., Waanders, G., Ferrero, I., Anjuère, F., Acha-Orbea, H., & MacDonald, H. R. (1996). Expression and presentation of endogenous mouse mammary tumor virus superantigens by thymic and splenic dendritic cells and B cells. *The Journal of Immunology*, *157*(7), 2789-2794.
38. Huber, B. T. (1995). The role of superantigens in virus infection. *Journal of Clinical Immunology*, *15*(6), S22-S25.
39. Pullen, A. M., Wade, T., Marrack, P., & Kappler, J. W. (1990). Identification of the region of T cell receptor β chain that interacts with the self-superantigen MIs-1a. *Cell*, *61*(7), 1365-1374.
40. Liu, J., Bramblett, D., Zhu, Q., Lozano, M., Kobayashi, R., Ross, S. R., & Dudley, J. P. (1997). The matrix attachment region-binding protein SATB1 participates in negative regulation of tissue-specific gene expression. *Molecular and Cellular Biology*, *17*(9), 5275-5287.

41. Maitra, U., Seo, J., Lozano, M. M., & Dudley, J. P. (2006). Differentiation-induced cleavage of Cutl1/CDP generates a novel dominant-negative isoform that regulates mammary gene expression. *Molecular and Cellular Biology*, *26*(20), 7466-7478.
42. Callahan, R., & Smith, G. H. (2008). Common integration sites for MMTV in viral induced mouse mammary tumors. *Journal of Mammary Gland Biology and Neoplasia*, *13*(3), 309-321.
43. Cato, A. C., Henderson, D., & Ponta, H. (1987). The hormone response element of the mouse mammary tumor virus DNA mediates the progestin and androgen induction of transcription in the proviral long terminal repeat region. *The EMBO Journal*, *6*(2), 363-368.
44. Mok, E., Golovkina, T. V., & Ross, S. R. (1992). A mouse mammary tumor virus mammary gland enhancer confers tissue-specific but not lactation-dependent expression in transgenic mice. *Journal of Virology*, *66*(12), 7529-7532.
45. Yanagawa, S., Tanaka, H., & Ishimoto, A. (1991). Identification of a novel mammary cell line-specific enhancer element in the long terminal repeat of mouse mammary tumor virus, which interacts with its hormone-responsive element. *Journal of Virology*, *65*(1), 526-531.
46. Ham, J., Thomson, A., Needham, M., Webb, P., & Parker, M. (1988). Characterization of response elements for androgens, glucocorticoids and progestins in mouse mammary tumour virus. *Nucleic Acids Research*, *16*(12), 5263-5276.
47. Nandi, S., & McGrath, C. M. (1973). Mammary neoplasia in mice. *Advances in Cancer Research* (Vol. 17, pp. 353-414). Academic Press.
48. Mertz, J. A., Lozano, M. M., & Dudley, J. P. (2009). Rev and Rex proteins of human complex retroviruses function with the MMTV Rem-responsive element. *Retrovirology*, *6*(1), 10.
49. Byun, H., Das, P., Yu, H., Aleman, A., Lozano, M. M., Matouschek, A., & Dudley, J. P. (2017). Mouse mammary tumor virus signal peptide uses a novel p97-dependent and derlin-independent retrotranslocation mechanism to escape proteasomal degradation. *MBio*, *8*(2), e00328-17.
50. Dultz, E., Hildenbeutel, M., Martoglio, B., Hochman, J., Dobberstein, B., & Kapp, K. (2008). The signal peptide of the mouse mammary tumor virus Rem protein is released from the endoplasmic reticulum membrane and accumulates in nucleoli. *Journal of Biological Chemistry*, *283*(15), 9966-9976.
51. Ball, J. K., & Dekaban, G. A. (1987). Characterization of early molecular biological events associated with thymic lymphoma induction following infection with a thymotropic type-B retrovirus. *Virology*, *161*(2), 357-365.
52. Ball, J. K., & McCarter, J. A. (1971). Repeated demonstration of a mouse leukemia virus after treatment with chemical carcinogens. *Journal of the National Cancer Institute*, *46*(4), 751-762.

53. Ball, J. K., Dekaban, G. A., McCarter, J. A., & Loosmore, S. M. (1983). Molecular biological characterization of a highly leukaemogenic virus isolated from the mouse. III. Identity with mouse mammary tumour virus. *Journal of General Virology*, *64*(10), 2177-2190.
54. Ball, J. K., Arthur, L. O., & Dekaban, G. A. (1985). The involvement of a type-B retrovirus in the induction of thymic lymphomas. *Virology*, *140*(1), 159-172.
55. Dekaban, G. A., & Ball, J. K. (1984). Integration of type B retroviral DNA in virus-induced primary murine thymic lymphomas. *Journal of Virology*, *52*(3), 784-792.
56. Ball, J. K., Diggelmann, H., Dekaban, G. A., Grossi, G. F., Semmler, R., Waight, P. A., & Fletcher, R. F. (1988). Alterations in the U3 region of the long terminal repeat of an infectious thymotropic type B retrovirus. *Journal of Virology*, *62*(8), 2985-2993.
57. Lee, J. W., Moffitt, P. G., Morley, K. L., & Peterson, D. O. (1991). Multipartite structure of a negative regulatory element associated with a steroid hormone-inducible promoter. *Journal of Biological Chemistry*, *266*(35), 24101-24108.
58. Liu, J., Barnett, A., Neufeld, E. J., & Dudley, J. P. (1999). Homeoproteins CDP and SATB1 interact: potential for tissue-specific regulation. *Molecular and Cellular Biology*, *19*(7), 4918-4926.
59. Morley, K. L., Toohey, M. G., & Peterson, D. O. (1987). Transcriptional repression of a hormone-responsive promoter. *Nucleic Acids Research*, *15*(17), 6973-6989.
60. Zhu, Q., & Dudley, J. P. (2002). CDP binding to multiple sites in the mouse mammary tumor virus long terminal repeat suppresses basal and glucocorticoid-induced transcription. *Journal of Virology*, *76*(5), 2168-2179.
61. Zhu, Q., Gregg, K., Lozano, M., Liu, J., & Dudley, J. P. (2000). CDP is a repressor of mouse mammary tumor virus expression in the mammary gland. *Journal of Virology*, *74*(14), 6348-6357.
62. Bramblett, D., Hsu, C. L., Lozano, M., Earnest, K., Fabritius, C., & Dudley, J. (1995). A redundant nuclear protein binding site contributes to negative regulation of the mouse mammary tumor virus long terminal repeat. *Journal of Virology*, *69*(12), 7868-7876.
63. Ross, S. R., Hsu, C. L., Choi, Y., Mok, E., & Dudley, J. P. (1990). Negative regulation in correct tissue-specific expression of mouse mammary tumor virus in transgenic mice. *Molecular and Cellular Biology*, *10*(11), 5822-5829.
64. Mertz, J. A., Mustafa, F., Meyers, S., & Dudley, J. P. (2001). Type B leukemogenic virus has a T-cell-specific enhancer that binds AML-1. *Journal of Virology*, *75*(5), 2174-2184.
65. Mertz, J. A., Kobayashi, R., & Dudley, J. P. (2007). ALY is a common coactivator of RUNX1 and c-Myb on the type B leukemogenic virus enhancer. *Journal of Virology*, *81*(7), 3503-3513.
66. Broussard, D. R., Mertz, J. A., Lozano, M., & Dudley, J. P. (2002). Selection for c-myc integration sites in polyclonal T-cell lymphomas. *Journal of Virology*, *76*(5), 2087-2099.

67. Broussard, D. R., Lozano, M. M., & Dudley, J. P. (2004). Rorγ (Rorc) is a common integration site in type B leukemogenic virus-induced T-cell lymphomas. *Journal of Virology*, *78*(9), 4943-4946.
68. Denis, F., Shoukry, N. H., Delcourt, M., Thibodeau, J., Labrecque, N., McGrath, H., & Sékaly, R. P. (2000). Alternative proteolytic processing of mouse mammary tumor virus superantigens. *Journal of Virology*, *74*(7), 3067-3073.
69. McMahon, C. W., Bogatzki, L. Y., & Pullen, A. M. (1997). Mouse mammary tumor virus superantigens require N-linked glycosylation for effective presentation to T cells. *Virology*, *228*(2), 161-170.
70. Mustafa, F., Bhadra, S., Johnston, D., Lozano, M., & Dudley, J. P. (2003). The type B leukemogenic virus truncated superantigen is dispensable for T-cell lymphomagenesis. *Journal of Virology*, *77*(6), 3866-3870.
71. Carter Jr, C. W. (1995). The nucleoside deaminases for cytidine and adenosine: structure, transition state stabilization, mechanism, and evolution. *Biochimie*, *77*(1-2), 92-98.
72. Harris, R. S., Petersen-Mahrt, S. K., & Neuberger, M. S. (2002). RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Molecular Cell*, *10*(5), 1247-1253.
73. Petersen-Mahrt, S. K., Harris, R. S., & Neuberger, M. S. (2002). AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature*, *418*(6893), 99.
74. Aydin, H., Taylor, M. W., & Lee, J. E. (2014). Structure-guided analysis of the human APOBEC3-HIV restrictome. *Structure*, *22*(5), 668-684.
75. Desimmie, B. A., Delviks-Frankenberrry, K. A., Burdick, R. C., Qi, D., Izumi, T., & Pathak, V. K. (2014). Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. *Journal of Molecular Biology*, *426*(6), 1220-1245.
76. Feng, Y., Baig, T. T., Love, R. P., & Chelico, L. (2014). Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. *Frontiers in Microbiology*, *5*, 450.
77. Imahashi, M., Nakashima, M., & Iwatani, Y. (2012). Antiviral mechanism and biochemical basis of the human APOBEC3 family. *Frontiers in Microbiology*, *3*, 250.
78. Malim, M. H., & Bieniasz, P. D. (2012). HIV restriction factors and mechanisms of evasion. *Cold Spring Harbor Perspectives in Medicine*, a006940.
79. Refsland, E. W., & Harris, R. S. (2013). The APOBEC3 family of retroelement restriction factors. *Intrinsic Immunity* (pp. 1-27). Springer, Berlin, Heidelberg.
80. Shandilya, S. M., Bohn, M. F., & Schiffer, C. A. (2014). A computational analysis of the structural determinants of APOBEC3's catalytic activity and vulnerability to HIV-1 Vif. *Virology*, *471*, 105-116.
81. Strebel, K. (2013). HIV accessory proteins versus host restriction factors. *Current Opinion in Virology*, *3*(6), 692-699.
82. Harjes, S., Solomon, W. C., Li, M., Chen, K. M., Harjes, E., Harris, R. S., & Matsuo, H. (2013). Impact of H216 on the DNA binding and catalytic activities of the HIV restriction factor APOBEC3G. *Journal of Virology*, JVI-03173.

83. Nabel, C. S., Lee, J. W., Wang, L. C., & Kohli, R. M. (2013). Nucleic acid determinants for selective deamination of DNA over RNA by activation-induced deaminase. *Proceedings of the National Academy of Sciences*, *110*(35), 14225-14230.
84. Navaratnam, N., & Sarwar, R. (2006). An overview of cytidine deaminases. *International Journal of Hematology*, *83*(3), 195.
85. Conticello, S. G., Thomas, C. J., Petersen-Mahrt, S. K., & Neuberger, M. S. (2004). Evolution of the AID/APOBEC family of polynucleotide (deoxy) cytidine deaminases. *Molecular Biology and Evolution*, *22*(2), 367-377.
86. Harris, R. S., & Liddament, M. T. (2004). Retroviral restriction by APOBEC proteins. *Nature Reviews Immunology*, *4*(11), 868.
87. LaRue, R. S., Andrésdóttir, V., Blanchard, Y., Conticello, S. G., Derse, D., Emerman, M., & Malik, H. S. (2009). Guidelines for naming nonprimate APOBEC3 genes and proteins. *Journal of Virology*, *83*(2), 494-497.
88. LaRue, R. S., Jónsson, S. R., Silverstein, K. A., Lajoie, M., Bertrand, D., El-Mabrouk, N., & Harris, R. S. (2008). The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC Molecular Biology*, *9*(1), 104.
89. Teng, B., Burant, C. F., & Davidson, N. O. (1993). Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science*, *260*(5115), 1816-1819.
90. Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J., & Scott, J. (1987). A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell*, *50*(6), 831-840.
91. Rogozin, I. B., Basu, M. K., Jordan, I. K., Pavlov, Y. I., & Koonin, E. V. (2005). APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy) cytidine deaminases predicted by computational analysis. *Cell Cycle*, *4*(9), 1281-1285.
92. Mikl, M. C., Watt, I. N., Lu, M., Reik, W., Davies, S. L., Neuberger, M. S., & Rada, C. (2005). Mice deficient in APOBEC2 and APOBEC3. *Molecular and Cellular Biology*, *25*(16), 7270-7277.
93. Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., & Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, *102*(5), 553-563.
94. Arakawa, H., Hauschild, J., & Buerstedde, J. M. (2002). Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science*, *295*(5558), 1301-1306.
95. Robbiani, D. F., & Nussenzweig, M. C. (2013). Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. *Annual Review of Pathology: Mechanisms of Disease*, *8*, 79-103.
96. Okazaki, I. M., Hiai, H., Kakazu, N., Yamada, S., Muramatsu, M., Kinoshita, K., & Honjo, T. (2003). Constitutive expression of AID leads to tumorigenesis. *Journal of Experimental Medicine*, *197*(9), 1173-1181.

97. Harris, R. S., & Dudley, J. P. (2015). APOBECs and virus restriction. *Virology*, *479*, 131-145.
98. Li, M. M., & Emerman, M. (2011). Polymorphism in human APOBEC3H affects a dominant phenotype of subcellular localization and antiviral activity. *Journal of Virology*, JVI-00624.
99. Vieira, V. C., & Soares, M. A. (2013). The role of cytidine deaminases on innate immune responses against human viral infections. *BioMed Research International*, *2013*.
100. Larson, E. D., & Maizels, N. (2004). Transcription-coupled mutagenesis by the DNA deaminase AID. *Genome Biology*, *5*(3), 211.
101. Orthwein, A., & Di Noia, J. M. (2012, August). Activation induced deaminase: how much and where? *Seminars in Immunology* (Vol. 24, No. 4, pp. 246-254). Academic Press.
102. Kumar, R., DiMenna, L., Chaudhuri, J., & Evans, T. (2014). Biological function of activation-induced cytidine deaminase (AID). *Biomedical Journal*, *37*(5).
103. Gazumyan, A., Bothmer, A., Klein, I. A., Nussenzweig, M. C., & McBride, K. M. (2012). Activation-induced cytidine deaminase in antibody diversification and chromosome translocation. *Advances in Cancer Research* (Vol. 113, pp. 167-190). Academic Press.
104. Rajewsky, K. (1996). Clonal selection and learning in the antibody system. *Nature*, *381*(6585), 751.
105. Di Noia, J. M., & Neuberger, M. S. (2007). Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.*, *76*, 1-22.
106. Teng, G., & Papavasiliou, F. N. (2007). Immunoglobulin somatic hypermutation. *Annu. Rev. Genet.*, *41*, 107-120.
107. Peled, J. U., Kuang, F. L., Iglesias-Ussel, M. D., Roa, S., Kalis, S. L., Goodman, M. F., & Scharff, M. D. (2008). The biochemistry of somatic hypermutation. *Annu. Rev. Immunol.*, *26*, 481-511.
108. Stavnezer, J., Guikema, J. E., & Schrader, C. E. (2008). Mechanism and regulation of class switch recombination. *Annu. Rev. Immunol.*, *26*, 261-292.
109. Bransteitter, R., Pham, P., Calabrese, P., & Goodman, M. F. (2004). Biochemical analysis of hyper-mutational targeting by wild type and mutant AID. *Journal of Biological Chemistry*.
110. Halemano, K., Guo, K., Heilman, K. J., Barrett, B. S., Smith, D. S., Hasenkrug, K. J., & Santiago, M. L. (2014). Immunoglobulin somatic hypermutation by APOBEC3/Rfv3 during retroviral infection. *Proceedings of the National Academy of Sciences*, *111*(21), 7759-7764.
111. Nussenzweig, M. C., & Alt, F. W. (2004). Antibody diversity: one enzyme to rule them all. *Nature Medicine*, *10*(12), 1304.
112. Tran, T. H., Nakata, M., Suzuki, K., Begum, N. A., Shinkura, R., Fagarasan, S.,& Nagaoka, H. (2010). B cell–specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. *Nature Immunology*, *11*(2), 148.

113. de Yébenes, V. G., Belver, L., Pisano, D. G., González, S., Villasante, A., Croce, C., & Ramiro, A. R. (2008). miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *Journal of Experimental Medicine*, *205*(10), 2199-2206.
114. Teng, G., Hakimpour, P., Landgraf, P., Rice, A., Tuschl, T., Casellas, R., & Papavasiliou, F. N. (2008). MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity*, *28*(5), 621-629.
115. Dorsett, Y., Robbiani, D. F., Jankovic, M., Reina-San-Martin, B., Eisenreich, T. R., & Nussenzweig, M. C. (2007). A role for AID in chromosome translocations between c-myc and the IgH variable region. *Journal of Experimental Medicine*, *204*(9), 2225-2232.
116. Liu, M., Duke, J. L., Richter, D. J., Vinuesa, C. G., Goodnow, C. C., Kleinstein, S. H., & Schatz, D. G. (2008). Two levels of protection for the B cell genome during somatic hypermutation. *Nature*, *451*(7180), 841.
117. Pasqualucci, L., Migliazza, A., Fracchiolla, N., William, C., Neri, A., Baldini, L., & Dalla-Favera, R. (1998). BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proceedings of the National Academy of Sciences*, *95*(20), 11816-11821.
118. Robbiani, D. F., Bunting, S., Feldhahn, N., Bothmer, A., Camps, J., Deroubaix, S., & Ried, T. (2009). AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. *Molecular Cell*, *36*(4), 631-641.
119. Shen, H. M., Peters, A., Baron, B., Zhu, X., & Storb, U. (1998). Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science*, *280*(5370), 1750-1752.
120. Yamane, A., Resch, W., Kuo, N., Kuchen, S., Li, Z., Sun, H. W., & Casellas, R. (2011). Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nature Immunology*, *12*(1), 62.
121. Maul, R. W., & Gearhart, P. J. (2010). AID and somatic hypermutation. *Advances in Immunology* (Vol. 105, pp. 159-191). Academic Press.
122. Lee-Theilen, M., & Chaudhuri, J. (2010). Walking the AID tightrope. *Nature Immunology*, *11*(2), 107.
123. Geisberger, R., Rada, C., & Neuberger, M. S. (2009). The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proceedings of the National Academy of Sciences*, *106*(16), 6736-6741.
124. Aoufouchi, S., Faili, A., Zober, C., D'Orlando, O., Weller, S., Weill, J. C., & Reynaud, C. A. (2008). Proteasomal degradation restricts the nuclear lifespan of AID. *Journal of Experimental Medicine*, *205*(6), 1357-1368.
125. Uchimura, Y., Barton, L. F., Rada, C., & Neuberger, M. S. (2011). REG-γ associates with and modulates the abundance of nuclear activation-induced deaminase. *Journal of Experimental Medicine*, *208*(12), 2385-2391.

126. Cheng, H. L., Vuong, B. Q., Basu, U., Franklin, A., Schwer, B., Astarita, J., & Chaudhuri, J. (2009). Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. *Proceedings of the National Academy of Sciences*, pnas-0812304106.
127. Kumar, R., DiMenna, L., Schrode, N., Liu, T. C., Franck, P., Muñoz-Descalzo, S., & Evans, T. (2013). AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature*, *500*(7460), 89.
128. Qin, H., Suzuki, K., Nakata, M., Chikuma, S., Izumi, N., Maruya, M., & Nagaoka, H. (2011). Activation-induced cytidine deaminase expression in CD4+ T cells is associated with a unique IL-10-producing subset that increases with age. *PloS One*, *6*(12), e29141.
129. Nassal, M. (2008). Hepatitis B viruses: reverse transcription a different way. *Virus Research*, *134*(1-2), 235-249.
130. Liang, G., Liu, G., Kitamura, K., Wang, Z., Chowdhury, S., Monjurul, A. M., & Muramatsu, M. (2015). TGF-β suppression of HBV RNA through AID-dependent recruitment of an RNA exosome complex. *PLoS Pathogens*, *11*(4), e1004780.
131. Gourzi, P., Leonova, T., & Papavasiliou, F. N. (2007). Viral induction of AID is independent of the interferon and the Toll-like receptor signaling pathways but requires NF-κB. *Journal of Experimental Medicine*, *204*(2), 259-265.
132. MacDuff, D. A., Demorest, Z. L., & Harris, R. S. (2009). AID can restrict L1 retrotransposition suggesting a dual role in innate and adaptive immunity. *Nucleic Acids Research*, *37*(6), 1854-1867.
133. Tobollik, S., Meyer, L., Buettner, M., Klemmer, S., Kempkes, B., Kremmer, E., & Jungnickel, B. (2006). Epstein-Barr virus nuclear antigen 2 inhibits AID expression during EBV-driven B-cell growth. *Blood*, *108*(12), 3859-3864.
134. Bekerman, E., Jeon, D., Ardolino, M., & Coscoy, L. (2013). A role for host activation-induced cytidine deaminase in innate immune defense against KSHV. *PLoS Pathogens*, *9*(11), e1003748.
135. Apolonia, L., Schulz, R., Curk, T., Rocha, P., Swanson, C. M., Schaller, T., & Malim, M. H. (2015). Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. *PLoS Pathogens*, *11*(1), e1004609.
136. Bogerd, H. P., & Cullen, B. R. (2008). Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation. *RNA*, *14*(6), 1228-1236.
137. Zhen, A., Du, J., Zhou, X., Xiong, Y., & Yu, X. F. (2012). Reduced APOBEC3H variant anti-viral activities are associated with altered RNA binding activities. *PloS One*, *7*(7), e38771.
138. Hultquist, J. F., Lengyel, J. A., Refsland, E. W., LaRue, R. S., Lackey, L., Brown, W. L., & Harris, R. S. (2011). Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *Journal of Virology*, JVI-05238.

139. Refsland, E. W., Hultquist, J. F., Luengas, E. M., Ikeda, T., Shaban, N. M., Law, E. K., & Harris, R. S. (2014). Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity. *PLoS Genetics*, *10*(11), e1004761.
140. Refsland, E. W., Hultquist, J. F., & Harris, R. S. (2012). Endogenous origins of HIV-1 G-to-A hypermutation and restriction in the nonpermissive T cell line CEM2n. *PLoS Pathogens*, *8*(7), e1002800.
141. Yu, Q., König, R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., & Landau, N. R. (2004). Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nature Structural and Molecular Biology*, *11*(5), 435.
142. Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., & Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*, *424*(6944), 99.
143. Harris, R. S., Bishop, K. N., Sheehy, A. M., Craig, H. M., Petersen-Mahrt, S. K., Watt, I. N., & Malim, M. H. (2003). DNA deamination mediates innate immunity to retroviral infection. *Cell*, *113*(6), 803-809.
144. Liddament, M. T., Brown, W. L., Schumacher, A. J., & Harris, R. S. (2004). APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. *Current Biology*, *14*(15), 1385-1391.
145. Zhang, H., Yang, B., Pomerantz, R. J., Zhang, C., Arunachalam, S. C., & Gao, L. (2003). The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature*, *424*(6944), 94.
146. Iwatani, Y., Chan, D. S., Wang, F., Maynard, K. S., Sugiura, W., Gronenborn, A. M., & Levin, J. G. (2007). Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Research*, *35*(21), 7096-7108.
147. Newman, E. N., Holmes, R. K., Craig, H. M., Klein, K. C., Lingappa, J. R., Malim, M. H., & Sheehy, A. M. (2005). Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Current Biology*, *15*(2), 166-170.
148. Gillick, K., Pollpeter, D., Phalora, P., Kim, E. Y., Wolinsky, S. M., & Malim, M. H. (2013). Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4+ T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination. *Journal of Virology*, *87*(3), 1508-1517.
149. Albin, J. S., Brown, W. L., & Harris, R. S. (2014). Catalytic activity of APOBEC3F is required for efficient restriction of Vif-deficient human immunodeficiency virus. *Virology*, *450*, 49-54.
150. Cul5, T., Koizumi, Y., Takeuchi, J. S., Misawa, N., Kimura, Y., Morita, S., & Sato, K. (2014). Quantification of deaminase activity-dependent and-independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. *Journal of Virology*, JVI-00062.

151. Conticello, S. G., Harris, R. S., & Neuberger, M. S. (2003). The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Current Biology*, *13*(22), 2009-2013.
152. Zhang, W., Du, J., Evans, S. L., Yu, Y., & Yu, X. F. (2012). T-cell differentiation factor CBF-β regulates HIV-1 Vif-mediated evasion of host restriction. *Nature*, *481*(7381), 376.
153. Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., & Yu, X. F. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science*, *302*(5647), 1056-1060.
154. LaRue, R. S., Lengyel, J., Jónsson, S. R., Andrésdóttir, V., & Harris, R. S. (2010). Lentiviral Vif degrades the APOBEC3Z3/APOBEC3H protein of its mammalian host and is capable of cross-species activity. *Journal of Virology*, *84*(16), 8193-8201.
155. Han, X., Liang, W., Hua, D., Zhou, X., Du, J., Evans, S. L., & Zhang, W. (2014). Evolutionarily conserved requirement for core binding factor beta in the assembly of the human immunodeficiency virus/simian immunodeficiency virus Vif-cullin 5-RING E3 ubiquitin ligase. *Journal of Virology*, *88*(6), 3320-3328.
156. Zhang, W., Wang, H., Li, Z., Liu, X., Liu, G., Harris, R. S., & Yu, X. F. (2014). Cellular requirements for BIV Vif-mediated inactivation of bovine APOBEC3 proteins. *Journal of Virology*, JVI-02072.
157. Derse, D., Hill, S. A., Princler, G., Lloyd, P., & Heidecker, G. (2007). Resistance of human T cell leukemia virus type 1 to APOBEC3G restriction is mediated by elements in nucleocapsid. *Proceedings of the National Academy of Sciences*, *104*(8), 2915-2920.
158. Mahieux, R., Suspene, R., Delebecque, F., Henry, M., Schwartz, O., Wain-Hobson, S., & Vartanian, J. P. (2005). Extensive editing of a small fraction of human T-cell leukemia virus type 1 genomes by four APOBEC3 cytidine deaminases. *Journal of General Virology*, *86*(9), 2489-2494.
159. Ooms, M., Krikoni, A., Kress, A. K., Simon, V., & Münk, C. (2012). APOBEC3A, APOBEC3B and APOBEC3H haplotype 2 restrict human T-lymphotropic virus type I (HTLV-1). *Journal of Virology*, JVI-06570.
160. Sasada, A., Takaori-Kondo, A., Shirakawa, K., Kobayashi, M., Abudu, A., Hishizawa, M., & Uchiyama, T. (2005). APOBEC3G targets human T-cell leukemia virus type 1. *Retrovirology*, *2*(1), 32.
161. Strebel, K. (2005). APOBEC3G & HTLV-1: inhibition without deamination. *Retrovirology*, *2*(1), 37.
162. Stenglein, M. D., & Harris, R. S. (2006). APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *Journal of Biological Chemistry*, *281*(25), 16837-16841.
163. Bogerd, H. P., Wiegand, H. L., Hulme, A. E., Garcia-Perez, J. L., O'Shea, K. S., Moran, J. V., & Cullen, B. R. (2006). Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proceedings of the National Academy of Sciences*, *103*(23), 8780-8785.

164. Chiu, Y. L., Witkowska, H. E., Hall, S. C., Santiago, M., Soros, V. B., Esnault, C., & Greene, W. C. (2006). High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proceedings of the National Academy of Sciences*, *103*(42), 15588-15593.
165. Burns, M. B., Lackey, L., Carpenter, M. A., Rathore, A., Land, A. M., Leonard, B., & Yee, D. (2013). APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature*, *494*(7437), 366.
166. Bogerd, H. P., Wiegand, H. L., Doehle, B. P., Lueders, K. K., & Cullen, B. R. (2006). APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Research*, *34*(1), 89-95.
167. Takeda, E., Tsuji-Kawahara, S., Sakamoto, M., Langlois, M. A., Neuberger, M. S., Rada, C., & Miyazawa, M. (2008). Mouse APOBEC3 restricts friend leukemia virus infection and pathogenesis in vivo. *Journal of Virology*, *82*(22), 10998-11008.
168. Santiago, M. L., Montano, M., Benitez, R., Messer, R. J., Yonemoto, W., Chesebro, B., & Greene, W. C. (2008). Apobec3 encodes Rfv3, a gene influencing neutralizing antibody control of retrovirus infection. *Science*, *321*(5894), 1343-1346.
169. Low, A., Okeoma, C. M., Lovsin, N., de las Heras, M., Taylor, T. H., Peterlin, B. M., & Fan, H. (2009). Enhanced replication and pathogenesis of Moloney murine leukemia virus in mice defective in the murine APOBEC3 gene. *Virology*, *385*(2), 455-463.
170. Stavrou, S., Nitta, T., Kotla, S., Ha, D., Nagashima, K., Rein, A. R., & Ross, S. R. (2013). Murine leukemia virus glycosylated Gag blocks apolipoprotein B editing complex 3 and cytosolic sensor access to the reverse transcription complex. *Proceedings of the National Academy of Sciences*, *110*(22), 9078-9083.
171. Langlois, M. A., Kemmerich, K., Rada, C., & Neuberger, M. S. (2009). The AKV murine leukemia virus is restricted and hypermutated by mouse APOBEC3. *Journal of Virology*, *83*(22), 11550-11559.
172. Browne, E. P., & Littman, D. R. (2008). Species-specific restriction of apobec3-mediated hypermutation. *Journal of Virology*, *82*(3), 1305-1313.
173. Rulli, S. J., Mirro, J., Hill, S. A., Lloyd, P., Gorelick, R. J., Coffin, J. M., & Rein, A. (2008). Interactions of murine APOBEC3 and human APOBEC3G with murine leukemia viruses. *Journal of Virology*, *82*(13), 6566-6575.
174. Mariani, R., Chen, D., Schröfelbauer, B., Navarro, F., König, R., Bollman, B., & Landau, N. R. (2003). Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell*, *114*(1), 21-31.
175. Doehle, B. P., Schäfer, A., Wiegand, H. L., Bogerd, H. P., & Cullen, B. R. (2005). Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. *Journal of Virology*, *79*(13), 8201-8207.
176. Abudu, A., Takaori-Kondo, A., Izumi, T., Shirakawa, K., Kobayashi, M., Sasada, A., & Uchiyama, T. (2006). Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. *Current Biology*, *16*(15), 1565-1570.

177. Jones, K. S., Ruscetti, S., & Lilly, F. (1988). Loss of pathogenicity of spleen focus-forming virus after pseudotyping with Akv. *Journal of Virology*, *62*(2), 511-518.
178. Bernstein, A., Mak, T. W., & Stephenson, J. R. (1977). The Friend virus genome: evidence for the stable association of MuLV sequences and sequences involved in erythroleukemic transformation. *Cell*, *12*(1), 287-294.
179. Okeoma, C. M., Lovsin, N., Peterlin, B. M., & Ross, S. R. (2007). APOBEC3 inhibits mouse mammary tumour virus replication in vivo. *Nature*, *445*(7130), 927.
180. MacMillan, A. L., Kohli, R. M., & Ross, S. R. (2013). APOBEC3 inhibition of MMTV infection: The role of cytidine deamination versus inhibition of reverse transcription. *Journal of Virology*, JVI-00112.
181. Nair, S., Sanchez-Martinez, S., Ji, X., & Rein, A. (2014). Biochemical and biological studies on mouse APOBEC3. *Journal of Virology*, JVI-03456.
182. Okeoma, C. M., Huegel, A. L., Lingappa, J., Feldman, M. D., & Ross, S. R. (2010). APOBEC3 proteins expressed in mammary epithelial cells are packaged into retroviruses and can restrict transmission of milk-borne virions. *Cell Host & Microbe*, *8*(6), 534-543.
183. Abudu, A., Takaori-Kondo, A., Izumi, T., Shirakawa, K., Kobayashi, M., Sasada, A., & Uchiyama, T. (2006). Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. *Current Biology*, *16*(15), 1565-1570.
184. Li, J., Hakata, Y., Takeda, E., Liu, Q., Iwatani, Y., Kozak, C. A., & Miyazawa, M. (2012). Two genetic determinants acquired late in Mus evolution regulate the inclusion of exon 5, which alters mouse APOBEC3 translation efficiency. *PLoS Pathogens*, *8*(1), e1002478.
185. Mustafa, F., Lozano, M., & Dudley, J. P. (2000). C3H mouse mammary tumor virus superantigen function requires a splice donor site in the envelope gene. *Journal of Virology*, *74*(20), 9431-9440.
186. Shackleford, G. M., & Varmus, H. E. (1988). Construction of a clonable, infectious, and tumorigenic mouse mammary tumor virus provirus and a derivative genetic vector. *Proceedings of the National Academy of Sciences*, *85*(24), 9655-9659.
187. Rada, C., Jarvis, J. M., & Milstein, C. (2002). AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proceedings of the National Academy of Sciences*, *99*(10), 7003-7008.
188. Ramiro, A. R., Jankovic, M., Callen, E., Difilippantonio, S., Chen, H. T., McBride, K. M., & Nussenzweig, A. (2006). Role of genomic instability and p53 in AID-induced c-myc–Igh translocations. *Nature*, *440*(7080), 105.
189. Mustafa, F., Al Amri, D., Al Ali, F., Al Sari, N., Al Suwaidi, S., Jayanth, P., & Rizvi, T. A. (2012). Sequences within both the 5′ UTR and Gag are required for optimal in vivo packaging and propagation of mouse mammary tumor virus (MMTV) genomic RNA. *PLoS One*, *7*(10), e47088.
190. Geiselhart, V., Schwantes, A., Bastone, P., Frech, M., & Löchelt, M. (2003). Features of the Env leader protein and the N-terminal Gag domain of feline foamy virus important for virus morphogenesis. *Virology*, *310*(2), 235-244.

191. Löchelt, M., Romen, F., Bastone, P., Muckenfuss, H., Kirchner, N., Kim, Y. B., & Flory, E. (2005). The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. *Proceedings of the National Academy of Sciences*, *102*(22), 7982-7987.
192. Purdy, A., Case, L., Duvall, M., Overstrom-Coleman, M., Monnier, N., Chervonsky, A., & Golovkina, T. (2003). Unique resistance of I/LnJ mice to a retrovirus is due to sustained interferon γ–dependent production of virus-neutralizing antibodies. *Journal of Experimental Medicine*, *197*(2), 233-243.
193. Beutner, U., Kraus, E., Kitamura, D., Rajewsky, K., & Huber, B. T. (1994). B cells are essential for murine mammary tumor virus transmission, but not for presentation of endogenous superantigens. *Journal of Experimental Medicine*, *179*(5), 1457-1466.
194. Gerpe, M. C. R., Renner, T. M., Bélanger, K., Lam, C., Aydin, H., & Langlois, M. A. (2015). N-linked glycosylation protects gammaretroviruses against deamination by APOBEC3 proteins. *Journal of Virology*, *89*(4), 2342-2357.
195. Kolokithas, A., Rosenke, K., Malik, F., Hendrick, D., Swanson, L., Santiago, M. L., & Evans, L. H. (2010). The glycosylated Gag protein of a murine leukemia virus inhibits the antiretroviral function of APOBEC3. *Journal of Virology*, *84*(20), 10933-10936.
196. Sheehy, A. M., Gaddis, N. C., & Malim, M. H. (2003). The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature Medicine*, *9*(11), 1404.
197. Stopak, K., de Noronha, C., Yonemoto, W., & Greene, W. C. (2003). HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Molecular Cell*, *12*(3), 591-601.
198. Marin, M., Rose, K. M., Kozak, S. L., & Kabat, D. (2003). HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nature Medicine*, *9*(11), 1398.
199. Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O., & Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *Journal of Biological Chemistry*, *274*(26), 18470-18476.
200. Wu, X., Feng, J., Komori, A., Kim, E. C., Zan, H., & Casali, P. (2003). Immunoglobulin somatic hypermutation: double-strand DNA breaks, AID and error-prone DNA repair. *Journal of Clinical Immunology*, *23*(4), 235-246.
201. Zan, H., Wu, X., Komori, A., Holloman, W. K., & Casali, P. (2003). AID-dependent generation of resected double-strand DNA breaks and recruitment of Rad52/Rad51 in somatic hypermutation. *Immunity*, *18*(6), 727-738.
202. Yu, K., Huang, F. T., & Lieber, M. R. (2004). DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. *Journal of Biological Chemistry*, *279*(8), 6496-6500.

203. Rogozin, I. B., & Diaz, M. (2004). Cutting edge: DGYW/WRCH is a better predictor of mutability at G: C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. *The Journal of Immunology*, *172*(6), 3382-3384.
204. Wei, L., Chahwan, R., Wang, S., Wang, X., Pham, P. T., Goodman, M. F., & MacCarthy, T. (2015). Overlapping hotspots in CDRs are critical sites for V region diversification. *Proceedings of the National Academy of Sciences*, *112*(7), E728-E737.
205. Golovkina, T. V., Dudley, J. P., Jaffe, A. B., & Ross, S. R. (1995). Mouse mammary tumor viruses with functional superantigen genes are selected during in vivo infection. *Proceedings of the National Academy of Sciences*, *92*(11), 4828-4832.
206. Dudley, J. P., Golovkina, T. V., & Ross, S. R. (2016). Lessons learned from mouse mammary tumor virus in animal models. *ILAR Journal*, *57*(1), 12-23.
207. Kobayashi, M., Takaori-Kondo, A., Miyauchi, Y., Iwai, K., & Uchiyama, T. (2005). Ubiquitination of APOBEC3G by an HIV-1 Vif-Cullin5-Elongin B-Elongin C complex is essential for Vif function. *Journal of Biological Chemistry*, *280*(19), 18573-18578.
208. Qin, W., Golovkina, T. V., Peng, T., Nepomnaschy, I., Buggiano, V., Piazzon, I., & Ross, S. R. (1999). Mammary gland expression of mouse mammary tumor virus is regulated by a novel element in the long terminal repeat. *Journal of Virology*, *73*(1), 368-376.
209. Miller, C. L., Garner, R., & Paetkau, V. (1992). An activation-dependent, T-lymphocyte-specific transcriptional activator in the mouse mammary tumor virus env gene. *Molecular and Cellular Biology*, *12*(7), 3262-3272.
210. Ichikawa, H. T., Sowden, M. P., Torelli, A. T., Bachl, J., Huang, P., Dance, G. S.,& Bottaro, A. (2006). Structural phylogenetic analysis of activation-induced deaminase function. *The Journal of Immunology*, *177*(1), 355-361.
211. Shinkura, R., Ito, S., Begum, N. A., Nagaoka, H., Muramatsu, M., Kinoshita, K., & Honjo, T. (2004). Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nature Immunology*, *5*(7), 707.
212. Honjo, T., Kobayashi, M., Begum, N., Kotani, A., Sabouri, S., & Nagaoka, H. (2012). The AID dilemma: infection, or cancer? *Advances in Cancer Research* (Vol. 113, pp. 1-44). Academic Press.
213. Okazaki, I. M., Kotani, A., & Honjo, T. (2007). Role of AID in tumorigenesis. *Advances in Immunology*, *94*, 245-273.
214. Lu, Z., Tsai, A. G., Akasaka, T., Ohno, H., Jiang, Y., Melnick, A. M., & Lieber, M. R. (2013). BCL6 breaks occur at different AID sequence motifs in Ig–BCL6 and non-Ig–BCL6 rearrangements. *Blood*, blood-2012.
215. Simpson, E., Dyson, P. J., Knight, A. M., Robinson, P. J., Elliott, J. I., & Altmann, D. M. (1993). T-cell receptor repertoire selection by mouse mammary tumor viruses and MHC molecules. *Immunological Reviews*, *131*(1), 93-115.

216. Han, J. H., Akira, S., Calame, K., Beutler, B., Selsing, E., & Imanishi-Kari, T. (2007). Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. *Immunity*, *27*(1), 64-75.
217. Cantaert, T., Schickel, J. N., Bannock, J. M., Ng, Y. S., Massad, C., Oe, T., & Al-Herz, W. (2015). Activation-induced cytidine deaminase expression in human B cell precursors is essential for central B cell tolerance. *Immunity*, *43*(5), 884-895.
218. Crouch, E. E., Li, Z., Takizawa, M., Fichtner-Feigl, S., Gourzi, P., Montaño, C., & Casellas, R. (2007). Regulation of AID expression in the immune response. *Journal of Experimental Medicine*, *204*(5), 1145-1156.
219. Watanabe, H., Numata, K., Ito, T., Takagi, K., & Matsukawa, A. (2004). Innate immune response in Th1-and Th2-dominant mouse strains. *Shock*, *22*(5), 460-466.
220. Kitamura, D., Roes, J., Kühn, R., & Rajewsky, K. (1991). A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin μ chain gene. *Nature*, *350*(6317), 423.
221. Gardinassi, L. G., & de Miranda Santos, I. K. F. (2016). Comment on "Regulation of immunity during visceral Leishmania infection" and further discussions about the role of antibodies in infections with Leishmania. *Parasites & Vectors*, *9*(1), 386.
222. Metzner, M., Jäck, H. M., & Wabl, M. (2012). LINE-1 retroelements complexed and inhibited by activation induced cytidine deaminase. *PLOS One*, *7*(11), e49358.
223. Chiu, Y. L., Witkowska, H. E., Hall, S. C., Santiago, M., Soros, V. B., Esnault, C., & Greene, W. C. (2006). High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proceedings of the National Academy of Sciences*, *103*(42), 15588-15593.

Vita

Gurvani Bhupindersingh was born in Dehradun, India on October 12, 1987, the daughter of Komal Singh and Bhupinder Singh. After completing her high school education in 2005, she studied Biotechnology at the University of Mumbai, India, graduating with a Bachelor of Engineering degree in 2009. Thereafter, in 2010 she received her Master of Science degree in Biotechnology from Texas A&M University, USA. She joined the Institute for Cell and Molecular Biology Ph.D. program at The University of Texas at Austin and started her work in Dr. Jaquelin Dudley's lab in June, 2013.

email: gur.vani@gmail.com

This dissertation was typed by the author.